Metal Ion Binding to Group II Intron Ribozymes: A Study from Bulk to the Single Molecule Level

DISSERTATION

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

MIRIAM STEINER

von

Biberist und Flumenthal SO

Promotionskomitee

Prof. Dr. Roland K. O. Sigel (Vorsitz, Leitung der Dissertation)

Prof. Dr. David Rueda

Prof. Dr. Heinz G. Berke

Zürich 2008

Content

List of abbreviations

1x TE	10 mM TrisHCl, pH 7.5, 1 mM EDTA
2AP	2-aminopurine
Av.	*Azotobacter vinelandii*
D135	domains 1, 3 and 5
DTT	(2S, 3S)-1,4-bissulfanylbutane-2,3-diol, dithiothreitol
E. coli/Ec.	*Escherichia coli*
EDTA	ethylenediamine-N,N,N',N'-tetraacetic acid
EtBr	ethidium bromide
groEL	60 kDa heat shock protein (chaperonine)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MOPS	3-morpholinopropane-1-sulfonic acid
NTP	nucleoside 5'-triphosphate
PAGE	polyacrylamide gel electrophoresis
Pl.	*Pylaiella littoralis*
ROS	reactive oxygen species
Sc.	*Saccharomyces cerevisiae*
Sc.ai5γ	yeast mitochondrial group II selfsplicing intron in the cytochrome c oxidase subunit 1 (*Sc*.cox1/5 intron)
TCEP	tris(2-carboxyethyl) phosphine hydrochloride
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol

Abstract

This work combines three approaches to study metal ion interactions with ribonucleic acid (RNA). The model systems used are large catalytic RNAs known as *group II intron ribozymes*. First, metal ion binding sites were located on a ribozyme sequence. Second, the influence of binding of several metal ions on the local structure of a ribozyme subelement was studied. The highlight in this work is a colorful fluorescence study of the structure and dynamics of the full-length molecules at the single molecule fluorescence level. The results permit the characterization of distinct effects of individual metal ions on RNA and imply a new paradigm for large RNA folding.

Kurzbeschrieb

Diese Arbeit beschreibt drei Ansätze zur Untersuchung des Zusammenspiels von Metallionen und Ribonukleinsäure (deutsch: RNS, englisch: RNA). Die hierbei verwendeten Modellsysteme sind grosse katalytische RNAs bekannt unter dem Namen *Group II Intron Ribozyme*. Zuerst wurden Stellen in der RNA Sequenz bestimmt, die Metallionen an sich binden. Danach wurde der Einfluss der Bindung verschiedener Metallionen auf die lokale Struktur eines kleinen Teilstücks aus dem Ribozym analysiert. Das Highlight der Arbeit ist eine farbenfrohe Fluoreszenz-Studie der Ribozymstruktur und -dynamik an einzelnen RNA-Molekülen. Die Ergebnisse zeigen die Einflüsse individueller Metallionen auf das Ribozym und beschreiben einen neuen Faltungsmechanismus für grosse RNA Moleküle.

1. Introduction

1.1 RNA: Pieces of genetic information

Genetic information of every organism out of the three domains of life (bacteria, archae and eucaryotes)[1] is stored in the form of deoxyribonucleic acid (DNA) in the morphologic unit of life, the cell. There, DNA is carefully maintained and protected from potential damage and only comes in contact with selected molecules for specialized purposes, the most important of which are reproduction/growth and the realization of genetic information for proper function of the organisms. For replication, the DNA is duplicated in an enzymatically mediated process to equip the daughter cells with a full set of DNA molecules before cell division. For expression of genetic information, a complex regulatory system controls the read-out of specific information depending on the need of the cells.

The DNA consists of a multitude of fragments called genes that contain information for a specific cellular function. In a first step for expression of a gene, the DNA serves as a template for the synthesis of ribonucleic acid (RNA) by the enzyme RNA polymerase in a process called transcription. The RNA molecules generated during transcription accomplish a variety of purposes. In the literature, RNA was first described for its involvement in protein synthesis.[2] RNA was found together with proteins in particles at the site of protein synthesis and was later on reported to provide the sequence information for the ribosome to link amino acids in the correct order to form proteins during translation.[3] This information gave rise to a basic scheme for transfer of genetic information (Figure 1) that became known as the d*ogma of biochemistry*. Therein, DNA is considered to serve exclusively for storage of genetic

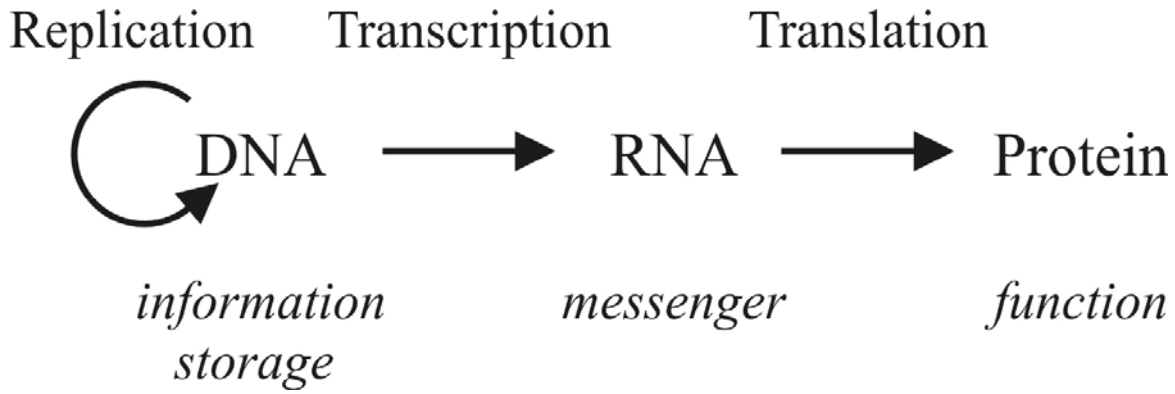


Figure 1. Transfer of genetic information – the dogma of biochemistry. DNA is duplicated before cell division to equip the daughter cells with a full set of DNA molecules, i.e. chromosomes. For expression of genetic information, DNA is transcribed into RNA, which is then translated into proteins. DNA serves for information storage, RNA represents the messenger and proteins implement function.

information, whereas the proteins provide some of the structural elements to build up the cell and represent the functional molecules that carry out the multiple enzymatic tasks needed in a living organism. RNA fills the role of intermediate messenger function that connects the clearly separated domains of storage (DNA) and structure/function (protein).

In reality, this simplified linear basic scheme is extended and refined by multiple processing activities at every step and every stage, e.g. RNA editing, capping and splicing, co- and posttranslational protein modification. Additionally, not every RNA molecule that is generated ultimately results in expression of a protein. RNA is also found to serve diverse structural or functional roles, often in complex with proteins that perform most elementary but very sophisticated functions, i.e. in the ribosome or the spliceosome. Therefore, RNA is found to be an important interposing key player along the way from storage to function and beyond. This work deals with a specific RNA molecule as an example of the fascinating complexity of processes developed in nature.

1.2 The RNA world hypothesis

DNA and proteins have been known and studied for quite some time, whereas the existence of RNA has only been approved a comparatively short time ago. RNA molecules are transcribed depending on the cell's needs at a certain point in time and disappear quickly upon degradation, e.g. after information transfer. This transient nature has made RNA difficult to detect, and the multiple functions of RNA are still being discovered as we go. Today, RNA molecules are known to act in various cellular processes like maintenance, transfer and processing of genetic information or cell development. RNA molecules are divided into classes according to their function. Some classes are described below:

- Messenger RNA (mRNA) and transfer RNA (tRNA) are involved in protein-synthesis.
- Ribosomal RNA (rRNA)[4] make up important components of the ribosome and therefore, the protein synthesis machinery of the cell.
- Small nuclear RNA (snRNA) are present in the spliceosome complex,[5] which plays an important role in RNA processing.
- Small RNAs (21-28 nts) like small interfering RNA (siRNA)[6, 7] take part in cell development through gene regulation.
- Micro RNA (miRNA)[8, 9] interact with and influence the translation and degradation of their target mRNA.

- The most recent finding identified the RNA group of riboswitches to be involved in gene regulation.[10, 11]

Apart from mRNA, all these RNA molecules are non-coding, meaning that no protein will ultimately origin from the RNA sequence.

In the early 80ties, pioneering studies described the ability of RNA molecules to provide catalytic activity in living cells without the help of amino acids.[12, 13] Before, protein enzymes were thought to be the sole catalytic workers to conduct the intermediate steps from DNA to proteins and for self-replication. Yet, RNA enzymes – in short ribozymes – occur among most of the RNA subgroups mentioned above. The natural catalytic reaction of ribozymes is a reversible phosphodiester transfer reaction of internucleotide bonds (hydrolysis or transesterification),[14, 15] which results in generation of two separate strand originating from a continuous polymer or vice versa. Additionally, ribozyme catalyzed reactions are also observed at carbon centers,[16, 17] i.e. peptide bond formation.[18] By precise positioning of nucleotides (nts) and possible additional components in the catalytic core, this reaction can be designed to be site-specific. *In vitro* selection and evolution have also generated ribozymes that are able to accelerate additional reactions, e.g. carbon-carbon bond formation by the Diels-Alder ribozyme,[19, 20] and further contribute to the vast catalytic repertoire of ribozyme reactions.

The ability of RNA to carry genetic information in viruses plus the finding that RNA itself is able to perform catalytic activity similar to enzymes has lead to the *RNA world hypothesis* in 1986, in which the earliest form of life relies on RNA only.[21] RNA combines both informational and catalytic properties for self-replication in a single molecular species. Therefore, RNA might reflect an ancestor molecule, from which DNA and proteins have evolved. In a first step, proteins were generated as more efficient and faster enzymes than the RNA counterparts. At a later state, DNA was generated by reverse transcription to represent a more stable form for storage of information than RNA. DNA is error-correcting because of its double-stranded structure but still capable of mutation and recombination. As information on the diverse functions of RNA is growing and further constitute pieces of the origin of life puzzle, this hypothesis has so far not been contradicted.[22]

1.3 RNA chemical structure and base pair formation

1.3.1 Nucleotides and polymerization

The basic components of RNA are sugar residues, phosphoryl groups and the four nucleic acids adenine (A), cytosine (C), guanine (G) and uracil (U). The ribose sugars are covalently linked to the nucleobases at their 1'-positions and carry a 5'-triphosphate group to make up the nucleotide triphosphate (NTP) building blocks ATP, CTP, GTP and UTP (Figure 2a). A RNA polymer of various sequence and length is synthesized by RNA polymerase using a DNA template and NTPs. The free 3'-OH at the end of a RNA strand attacks the α–P of the incoming NTP, resulting in attachment of the next base to the RNA polymer in a 3'-5' linkage. The cleavage of the triphosphate followed by release of pyrophosphate (Figure 2b) and its subsequent hydrolysis to inorganic phosphate during RNA synthesis energetically drives the polymerization reaction and renders it irreversible. The carbon atom numbering scheme in the base moiety (1'-5') imposes a direction on the RNA. The nucleotides of an RNA sequence are conventionally listed in the 5'-3' linkage direction and the ribose numbering scheme assigns the first 5'- and the last 3'-substitution of an RNA polymer to the 5'- and 3'-ends, respectively.

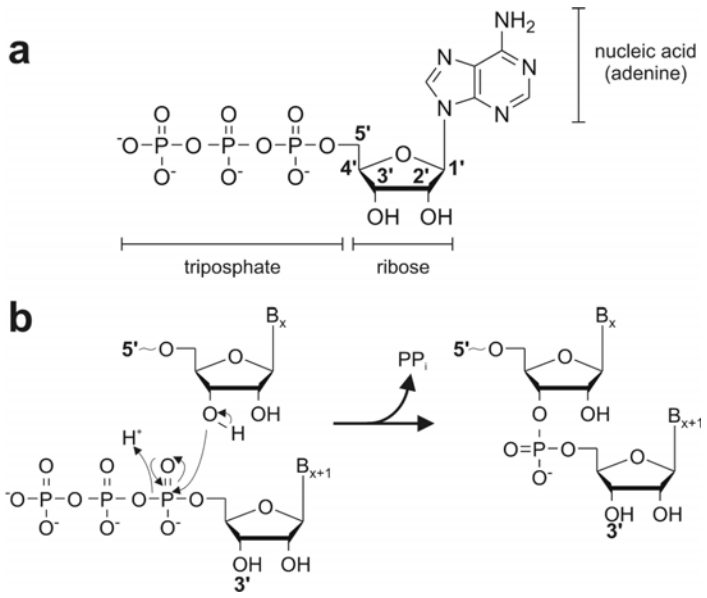


Figure 2. Nucleotide triphosphates (NTP) and polymerization. **a** Chemical depiction of adenosine triphospate (ATP) and its subunits. **b** Elongation of an RNA polymer by one nucleotide base (B). Polymerization reaction is catalyzed by RNA polymerase and releases pyrophosphate (PP_i).

1.3.2 Base pairing

Nucleotides are able to form specific pairs following a hydrogen bonding pattern that is dependent on the appropriate geometry of the hydrogen bond donors and acceptors in the bases. For the most common and stable basepairs, A pairs with U or thymine (T, in DNA), and G pairs with C in a canonical Watson-Crick (WC) type fashion (Figure 3a, b). In RNA, further base pairing schemes like the G●U wobble (Figure 3c) and (reverse) Hoogsteen base pairs are important for the formation of complex three dimensional structures, too. The

hydrogen bonding results in increased stabilization of the macromolecules and leads to the fact that nucleic acids often consist of two strands of paired nucleotides, in which the sequence of one strand implies the sequence of the other rendering the strands complementary. Formation of a daughter DNA molecule during replication and RNA formation during transcription relies on this complementarity and determines the nucleotide sequence in the newly synthesized strand. Furthermore, base pairing is the mechanism by which codons on mRNA are recognized by anticodons on tRNA during protein translation, and there are DNA- and RNA-binding proteins that specifically recognize base pairing patterns and identify specific regions of interests.

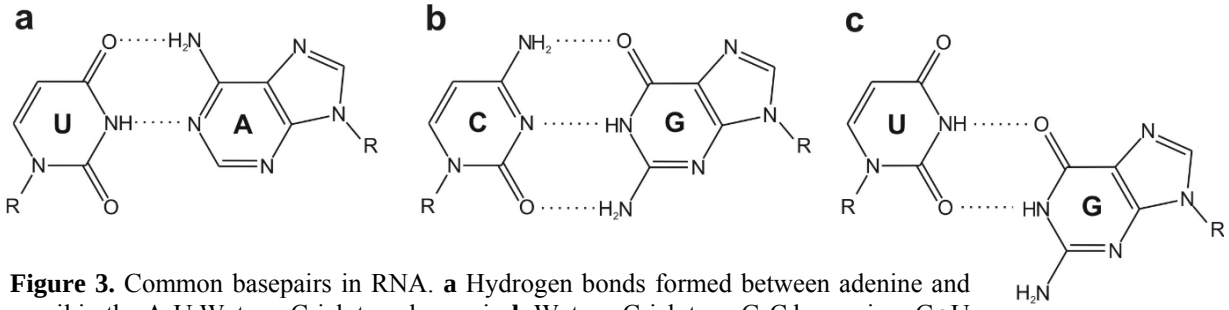


Figure 3. Common basepairs in RNA. **a** Hydrogen bonds formed between adenine and uracil in the A-U Watson-Crick type basepair. **b** Watson Crick type G-C basepair. **c** G●U wobble basepair.

1.3.3 Double helices in RNA and DNA

RNA is a structural analog of DNA, but differs from DNA in that

a) RNA uses an unmethylated form of thymine (T) called uracil (U) as a building block,
b) RNA contains an additional 2'-hydroxyl group on the ribose subunit (DNA: deoxyribonucleic acid, RNA: ribonucleic acid)
c) RNA molecules are usually much shorter in length than DNA and RNA is often in single stranded (ss) form for its biological roles, whereas DNA is double stranded (ds).

Double stranded RNA (dsRNA) regions generally adopt an A-form helix, whereas dsDNA exists in the well-known B-form helix described by Watson and Crick (Figure 4a, b).[23] The A-form helix of RNA (Figure 4c) has a diameter of about 26 Å and eleven basepairs per turn at a pitch of 30 Å.[24] This helix is characterized by a deep, narrow major groove and a shallow, wide minor groove. The shallow minor groove provides a suitable platform for interactions with various partners. Additionally, the nucleobase pairs are tilted by 14° normal to the helix axis of the double strand, which gives rise to the typical hole in the middle of the double helix when looking down the helix axis (Figure 4d). Due to the subtle differences mentioned above, RNA displays a more flexible molecule than DNA and is able

to adopt helical structures of sequence complementary stretches as well as single stranded three-dimensional architectural motifs like loops, turns and bulges, all within one single polymer chain.

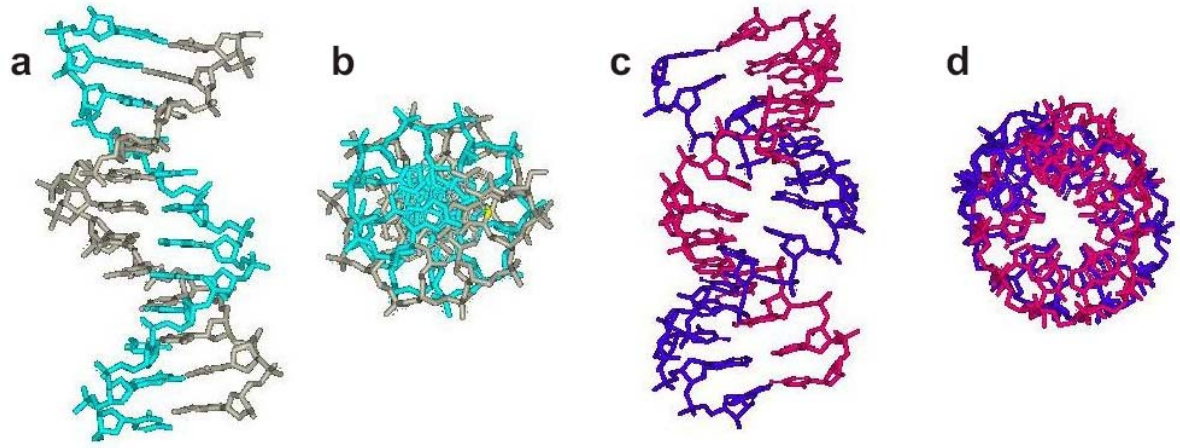


Figure 4. Helices formed by hydrogen bonding of ribonucleotide polymers. **a** B-form DNA helix. The two separate strands are shown in grey and light blue. **b** DNA helix from a, view from above. **c** double stranded RNA helix (A-form). The strands are shown in purple and pink. **d** RNA helix from c viewed from above. Figure is adapted from [25].

1.4 Metal ions in RNA folding and catalysis

The negatively charged phosphate-sugar backbone turns the RNA into a polyanion. Starting from this highly charged random coil structure, RNA forms the secondary as well as a tightly packed tertiary structure, in which negative charges accumulate in close space. Obviously, there is a need for a neutralizing agent with positive charges of high density to stabilize the structure of higher order. First and foremost, this task is met by metal cations.

The major task of electrostatic relaxation of the backbone occurs independently of the metal ion identity but tends to be taken over by monovalent metal ions, i.e. K^+ and Na^+.[26-28] These interactions are mainly unspecific and allow the formation of secondary structure. Tertiary compaction on the other hand generally involves essential interactions with specifically bound and dehydrated metal ions of higher valency, e.g. Mg^{2+}, to correctly orient the structural elements in space.[29, 30] Yet in some cases, specific monovalent ions can also be important for this task.[28, 31, 32] Furthermore, polycharged organic cations like spermine and RNA binding proteins can replace metal ions to form higher order structures.

In addition to the obvious implications in folding of RNA, metal ions have been described to be required for chemistry of many ribozymes. Divalent ions act as the most catalytic effective cofactors. These ions are found within the active site of ribozymes, where they directly take part in acid-base catalysis,[33-35] or generally facilitate formation of

transition states.[36] Therefore, these ribozyme are divalent metal ion dependent ribozymes. Yet this is not an universal concept, since the catalysis can be stimulated in the absence of Mg^{2+} in certain ribozymes.[37] In these cases, Mg^{2+} or other divalent metal ions are not specifically involved in the chemical mechanism of catalysis. Nevertheless, ribozymes are always associated with metal ions and therefore, metalloribozymes in a more general term.

The number of potential metal ion binding sites on RNA is large. All four nucleotides provide several donor atoms for metal ion coordination, i.e. the atoms of the phosphodiester linkage, the carbonyl oxygens and all the unprotonated nitrogen atoms in the bases.[38] At physiologic pH, these atoms are deprotonated and accessible for metal ion binding. The coordination sites in RNA show a lower affinity for metal ion binding compared to sites in proteins (sulfides in Cys, the imidazole moiety in His and carboxylates in Asp and Glu).[38] As a result, the metal ion binding in RNA is generally nonspecific and dynamic.[38] In addition, there are at least four ways for metal ions to bind to RNA (nonspecific charge screening ions, electrostatically localized ions, inner- or outer-sphere site-bound ions) and metal ions can only rarely be categorized into strictly one of these classes.[38-40] The identity of the metal ion that binds the RNA is of great importance and interest, too, since every metal ion behaves differently in terms of binding strength, coordination preferences and size.[41]

Taken together, RNA molecules are always associated with metal ions.[42-45] The common natural cofactors are K^+ and Mg^{2+}, since they are the most abundant metal ions in the cell. They serve to compensate negative charges in the backbone, promote structural stabilization and are frequently involved in the chemical mechanism of catalysis. The binding of mono- and divalent metal ions to nucleic acids for successful folding and catalysis is complex and only beginning to be understood as it is a subject of current research.[38, 44, 46] Divalent metal ions are often essential for folding and catalysis of ribozymes, although they can in some cases be replaced by high concentrations of monovalent metal ions.[47]

1.5 RNA folding

1.5.1 RNA folding motifs/elements

One basic architectural element in RNA structure is the A-form helix. Helical regions have a high thermodynamic stability that can be calculated with reasonable accuracy and is used in models for prediction of RNA structure. The formation of proximal, partially self-complementary regions provides the basis for the arrangement of the secondary structure and the partitioning of the molecules into domains. Once the helical regions (Figure 5a) are

defined, further structural elements become apparent at the secondary structure level. Such RNA motifs are terminal or hairpin loops (Figure 5b), i.e. tetraloops,[48] bulges (Figure 5c),[49] internal loops (Figure 5d),[50] and multihelix or junction loops (Figure 5e), i.e. pseudoknots,[51] just to mention a few. These elements might be stable on their own or interact with a corresponding element to form short- or long-range tertiary interactions that organize and define the overall RNA architecture and 3D structure.

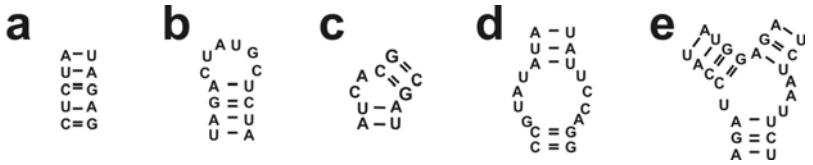


Figure 5. RNA motifs. **a** helical region, **b** terminal/hairpin loop (6 nt), **c** 2 nt bulge, **d** internal bulge, **e** multihelix junction.

1.5.2 RNA folding pathways

RNA molecules fold into a variety of complex shapes for their distinct biological activity. The nucleotide polymer chain not only includes the sequence information, but also defines the secondary, as well as the tertiary structure for proper function. The focus in RNA folding studies is on the mechanistic relation between the sequence and the functional, three dimensional structure. A molecule folds into its native 3D structure starting from random coil in a specific way within a short time of milliseconds to minutes rather than sampling all possible folded species. This is a process that would take far too long for a molecule designated to adapt to fast changing conditions for survival of an organism. This thought has been described by the Levinthal paradox for protein folding, and the considerations can also be adopted to RNA folding.[52] Generally, RNA folding is a hierarchical process, in which secondary elements form before the tertiary structure is built by formation of weak long-range interactions,[53] although later rearrangements of secondary structure elements are sometimes possible.[54] Generally, tertiary

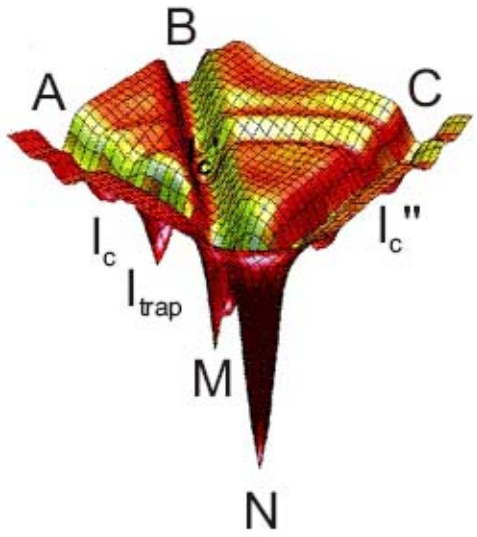


Figure 6. Section of funnel diagram showing RNA energy landscape and parallel folding pathways. Three possible folding channels A (low Na^+), B (intermediate Na^+) and C (high Na^+), starting from different points at the rim of the funnel, along with collapsed (I_c) and trapped (I_{trap}) intermediates, as well as misfolded (M) species and the native state (N) are shown. The vertical dimension indicates the free energy, and the relief is made up of internal free energy barriers. The figure is a schematic depiction of folding of the *Tetrahymena* ribozyme and adapted from [57].

structure formation is often preceded by an early compaction step.[55, 56] The characteristics of the folding pathway are specific for every individual molecule, depend on the conditions (ionic strength, temperature, cofactors) and give rise to a broad set of strikingly different folding mechanisms. The ultimate goal is to reach the native (active) state, which represents the structure with lowest free energy, in a relevant time range.

Folding pathways of RNA, and biomolecules in general, can be described energetically by a funnel diagram representing a multidimensional energy landscape, along the surface of which a molecule tumbles and travels (Figure 6).[57, 58] Driven by energy, folding molecules traverse "downhill" to the final folded state, which typically represents a narrow region at the center of the funnel. Intermediate structures might exist along the course and represent local energy minima. These intermediates can be stable for a certain period of time, but might be misfolded and reflect a kinetic trap. The escape from the intermediates requires an energy input and, in case of kinetic traps, usually depicts the rate-limiting step in folding.[59, 60] The landscape of RNA molecules consists of discrete folding pathways that are separated by large energy barriers. Initial conditions determine the starting state and give rise to parallel folding pathways. There are specific paths that are more advantageous than others, both in folding speed and energy consumption.

Formation of RNA tertiary structure is highly dependent on

i) RNA sequence and length,
ii) metal ion identity,
iii) metal ion concentration,
iv) temperature,
v) the presence of other factors like RNA binding proteins and polyamines.[61]
vi) chaperones
vii) helicases

Small RNAs usually fold in a fast and apparent two state process,[60] whereas larger RNAs generally follow a more rugged folding pathway.

1.6 Ribozymes

Ribozymes are suitable model systems to study the relationship of structure and function, since their catalytic activity directly reports on the extent of native structure formation. Eight structurally different main types of naturally occurring ribozymes have been identified, next to the autocatalytic RNA components that are part of ribonucleotide proteins (spliceosomal and ribosomal RNAs). According to their size and mechanism, the ribozymes are classified into the small or the large group (Table 1).

Table 1.

The naturally occurring ribozymes. The table is adapted from [81] and [94].

Ribozyme	Biological source	Size	Reaction products	Function (Reaction)	# identified
Small ribozymes					
Hammerhead ribozyme	Plant viroids and satellite RNAs	~40 nt	5'-OH 2',3'-cyclic phosphate	Replication (self-cleaving)	11
Hairpin ribozyme	Plant viroids and satellite RNAs	~60 nt	5'-OH 2',3'-cyclic phosphate	Replication (self-cleaving)	4
Hepatitis delta virus ribozyme (HDV)	Human hepatitis delta virus	~80 nt	5'-OH 2',3'-cyclic phosphate	Replication (self-cleaving)	2
VS ribozyme	*Neurospora crassa* mitochondria	154 nt	5'-OH 2',3'-cyclic phosphate	Replication (self-cleaving)	1
glmS ribozyme	Bacteria (gram positive)	~200 nt	5'-OH 2',3'-cyclic phosphate	Riboswitch (self-cleaving)	18
Large ribozymes					
RNase P RNA	Eucaryotes (nucleus and organelles) prokaryotes, bacteriophages	140-490 nt	Products with 5'-phosphate and 3'-OH	tRNA processing (Hydrolysis)	>300
Group I Introns	Eucaryotes (nucleus and mitochondria) prokaryotes, bacteriophages	200-1500 nt	Intron with 5'-guanosine and 3'-OH, 5'/3' ligated exons	Splicing (self-splicing)	>1000
Group II Introns	Eucaryotes (organelles) prokaryotes	300-3000 nt	Intron with 2'-5' lariat and 3'-OH, 5'/3' ligated exons	Splicing (self-splicing)	>700
further autocatalytic RNA components of RNPs					
Ribosomal RNAs	Eucaryotes, prokaryotes	5S-28S	Peptide bonds	Peptide bond formation (peptidyl transfer)	>5000
Spliceosomal RNAs	Eucaryotes	100-180 nt	Intron with 2'-5' lariat and 3'-OH, 5'/3' ligated exons	Splicing (trans-splicing)	>100

1.6.1 Small ribozymes

The best studied examples for naturally occurring small ribozymes are the hammerhead, the hairpin, the hepatitis delta virus (HDV) ribozyme and the Varkud Satellite (VS) ribozymes. Circular monomeric molecules of these RNA species (also multimeric in case of the VS ribozyme) are used as templates for a special type of replication known as the *rolling circle* mechanism, in which long linear transcripts are generated that consist of joined monomers. These RNA concatemers are then cut into monomers by catalytic motifs, which the monomers provide themselves. Therefore, these ribozymes are also termed *self-cleaving* RNAs. All small ribozymes use a common chemical reaction mechanism for cleavage of the phospodiester bond that generates a 2'-3'-cyclic phosphate and a product with a 5'-OH terminus (Figure 7). The monomers are subsequently recirculated for the next round of replication.

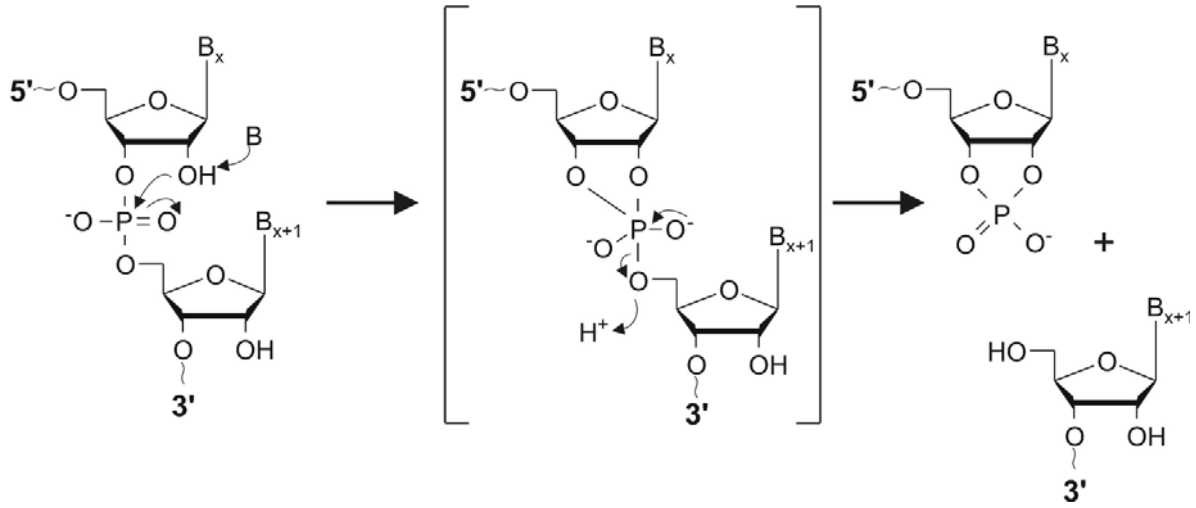


Figure 7. Mechanism of phosphodiester cleavage by the small ribozymes. The internal 2'-OH acts as the attacking nucleophile generating a 2'-3'-cyclic phosphate and a product with a 5'-OH terminus. B, base; Nu, nucleophile

All small ribozymes catalyze site-specific RNA backbone cleavage in a single stranded region of the molecule, but the way they achieve this reaction is very diverse. The ribozymes differ in sequence length, secondary and tertiary structure (Figure 8), as well as metal ion requirements for folding/catalysis. The hammerhead, the hairpin and the HDV ribozymes have a nucleotide sequence length of <85 nts, whereas the VS ribozyme is somewhat longer (160 nts) (Table 1). A wealth of biochemical and structural data has been gathered over the last few years and is reviewed in [15, 62-64]. (Partial) Crystal structures are available for all of the small ribozymes and helped to deduce the folding and cleavage mechanisms:[65-70]

- The hairpin ribozyme contains two independent folding domains, each with an internal loop and two flanking helices (Figure 8a). Initially, the folding units are coaxially stacked and rotate to form an antiparallel structure upon metal ion binding.[71, 72]
- The hammerhead ribozyme is built from three base-paired stem loops and a highly conserved core forming a three-way junction (Figure 8b). The hammerhead ribozyme folds in a defined two-stage process.[73] In the first stage, stems I and II form a coaxially stacked domain from an extended core structure. In the second step, stem III switches from being oriented in a similar direction to stem II to being closer to stem I.
- The HDV ribozyme is more complex but the structure is based on a pseudoknot with two stem-loops, in which the loop of the first stem-loop forms parts of the second stem (Figure 8c).[74] In contrast to the two ribozymes mentioned above, the HDV ribozyme does not undergo large conformational changes. Upon substrate binding, the molecule globally becomes only slightly more compact. Instead, it appears to be a local conformational change in and around the catalytic core that accompanies catalysis.[75]

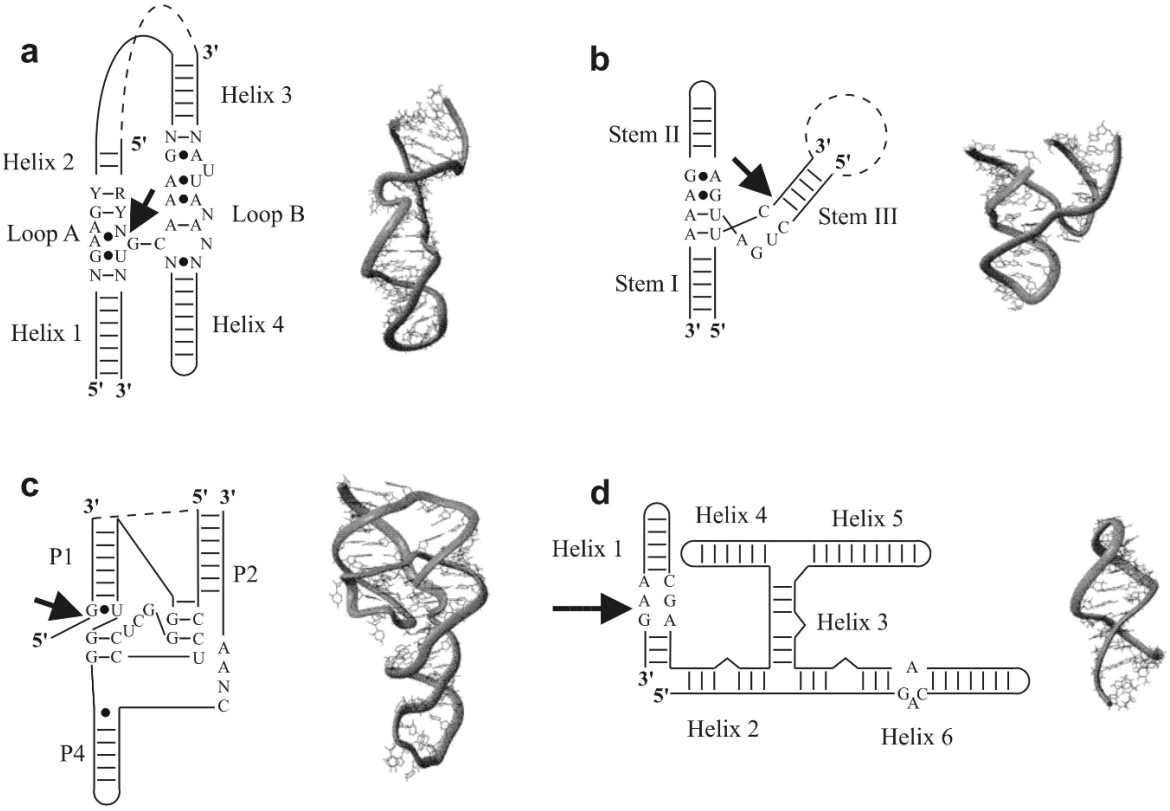


Figure 8. Structural diversity of small ribozymes: secondary and crystal structures. **a** Hairpin ribozyme. The crystal structure shown on the right depicts domain B (helices 3 and 4, PDB code 1B36) **b** Hammerhead ribozyme (488D) **c** Hepatitis delta virus (HDV) ribozyme (1CXO) **d** Varkud Satellite (VS) ribozyme. The crystal structure shows the cleavage site from the *Neurospora* VS ribozyme (1E4P). The arrows in the secondary structures indicate the site of cleavage. The figure is adapted from [76] and [81].

- The core of the VS ribozyme is made up of 6 helices, one substrate helix plus five core helices, and is organized in two three-way junctions (Figure 8d). Data on the VS ribozyme again implies a metal ion dependent two-stage folding of the junction formed by helices 2, 3, and 6.[62]

In 1998, the hammerhead, the hairpin and the VS ribozyme have been shown not to require Mg^{2+} for catalysis if monovalent metal ion like Li^{+} or Na^{+} are present in concentrations of 2-4 M.[37] Although the cleavage reaction was most efficient in the presence of divalent metal ions, a similar tertiary structure and cleavage products were found in high concentrations of monovalent metal ions. Therefore, divalent metal ions are not obligatory participants in the reaction and the tertiary structure contributes more to the catalytic reaction than previously recognized.

A further member of the class of small phosphoryl transfer ribozymes is the recently discovered *glms* ribozyme.[77] This RNA is the only natural occurring and self-cleaving ribozyme known to require a small molecule activator for catalysis, which is glucosamine-6-phosphate (GlcN6P).[78] It functions as a riboswitch, a molecule that changes its structure upon small molecule binding, and whose activator-dependent RNA cleavage regulates the *gmls* mRNA expression.

1.6.2 Large ribozymes

The group of the large ribozymes contains Ribonuclease P (RNase P), group I and group II introns (Table 1). These ribozymes have a sequence length of 100 to about 3000 nts (up to 6000 nts when counting the open reading frame (ORF) in group II introns - *vide infra*). RNase P is a site-specific endoribonuclease that processes precursor RNA substrate, i.e. the 5'-end of tRNAs[79, 80] or 5S rRNA, and it is mostly associated with proteins for its function. Therefore, RNase P is generally a ribonucleoprotein. Group I and II intron ribozymes belong to the classes of introns or intervening sequences. Intron sequences are non-coding parts of a gene, which interrupt the coding sequences (exons) and need to be removed for correct expression of a gene. This process happens at the RNA stage (Figure 1) and is called *splicing*. The group I and II introns are intronic RNA segments that cut themselves out of the larger RNA (exons plus introns) and hence, are called *self-splicing* introns. In distinction from *self-cleaving* introns, the *self-splicing* introns additionally rejoin the ends of two single cuts.[81]

The large ribozymes not only differ from small ribozymes in the catalytic activity, but also in their chemical reaction mechanism. In contrast to the small ribozymes that use an adjacent internal nucleophile for attack of the phosphodiester bond, the large ribozymes use

an external or far distant internal nucleophile. In case of RNase P, the attacking hydroxyl is a water molecule and the hydrolytic reaction results in products with a 5'-phosphate and a 3'-OH end. The catalytic activity of both the group I and group II introns consist of two transesterification steps. Group I introns use an external GTP (3'-OH of guanine) and group II introns the 2'-OH of an internal adenine as the nucleophile for the first step of splicing at the 5' splice site. In the second step for both introns, it is the 3'-OH of the newly released 5'-exon that attacks the 3'-splice site. As for the reaction of RNase P, both steps of splicing result in products with a 5'-phosphate and a 3'-OH end (Figure 9).[82, 83]

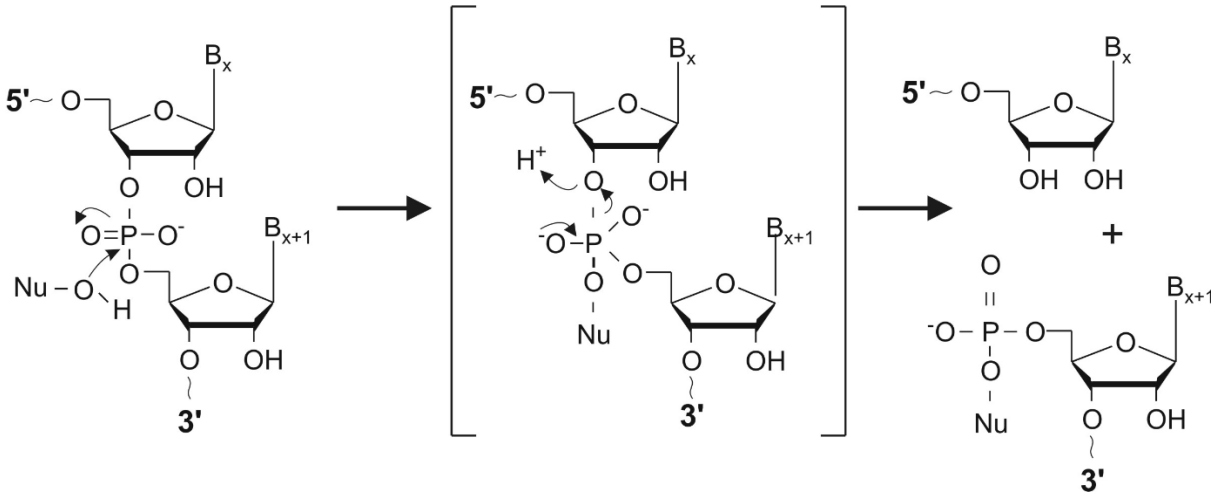


Figure 9. Mechanism of cleavage of phosphodiesterbonds by large ribozymes. The attacking nucleophile is a water molecule (RNase P), an external 3'-OH of guanine (group I introns, first step of splicing), an internal 2'-OH of adenine (group II introns, first step of splicing) or the 5'-OH of the 5'-exon (second step of splicing for both group I and II introns). The reactions produce a 5'-phosphate and a 3'-hydroxyl terminus. B, base; Nu, nucleophile.

The complexity of ribozymes increases with sequence length. All large ribozymes are multidomain molecules and the individual domains can be attributed to specific functions. The RNase P molecule is parted into a catalytic and a specificity domain (Figure 10a). Group I introns have a common core secondary structure of nine base-paired elements (P1-P9, Figure 10b).[15] The largest members are the group II introns, which share a typical secondary structure of six domains (Figure 10c). There are a number of models known for tertiary folding for large RNAs, all involving an initial stage of collapse followed by the formation of the native state in a second step.[27, 84] Folding rates for these states can differ considerably. The *Tetrahymena* group I intron collapses to a stable but misfolded species, which reflects a kinetic trap.[26, 85] In a slow second step, the molecule escapes from the trap through structural reorganization.[60, 86] Full-length RNase P behaves in a similar way, whereas its isolated catalytic domain shows a conformational search and collapse to a specific intermediate in a first fast step, before a local structural reorganization, i.e. consolidation around a metal ion binding site, takes place in a slower second step.[87, 88] Group II introns

behave differently, they are slow folders involving an obligate on-pathway intermediate (*vide infra*).

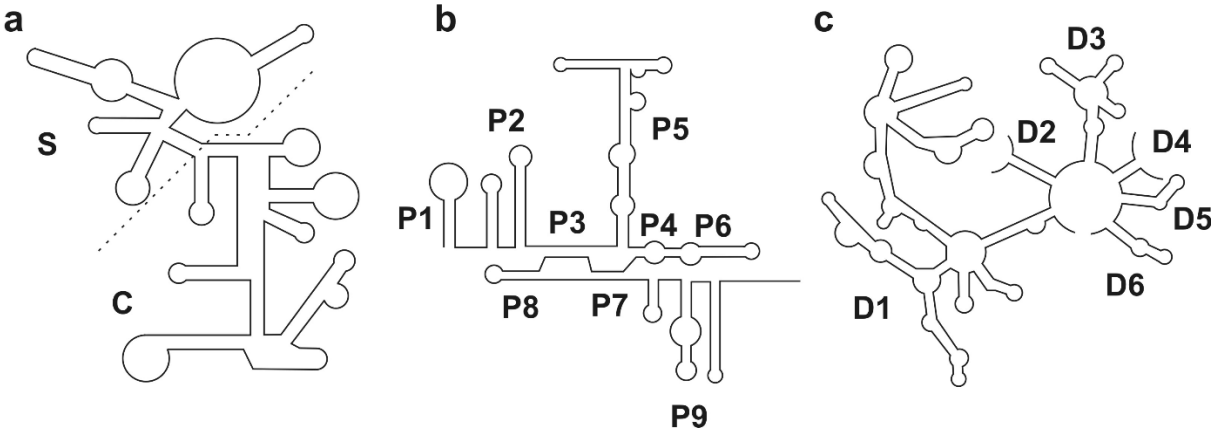


Figure 10. Schematic secondary structure depiction of the large ribozymes. **a** RNase P is divided (dashed line) into a specificity domain (S) and a catalytic domain (C). **b** Group I introns show a structure of nine base-paired elements P1-P9. **c** Group II introns comprise six domains (D1-6) that radiate from a central wheel. The schematic secondary structures shown are derived from *B. subtilis* (a), *T. thermophilus* (b) and *Sc*.ai5γ (c).

Crystallographic data on large ribozymes is so far only available on parts of the molecules. There are four recent crystal structures of group I intron ribozymes (reviewed in [89]) and there is evidence for a two metal ion mechanism in group I intron splicing.[90] All group I introns have been reported to essentially rely on divalent metal ions for folding and catalysis[91, 92] and a general two-metal ion cleavage mechanism has been stated for all large ribozymes.[93]

1.7 Group II intron ribozymes

1.7.1 The splicing reaction

The molecule studied in this work is a large RNA that belongs to the group II intron ribozymes. Group II introns are naturally occurring ribozymes ranging between 600 and several thousands nucleotides in size (Table 1). These RNAs are primarily found in organellar genes of bacteria and lower eucaryotes (mostly fungi and plants), but are absent from animal genomes.[94, 95] As introns, i.e. noncoding sequences on a pre-mRNA, they have to be removed from the primary transcript to yield a functional RNA. Group II intron ribozymes are remarkably large molecular machines, some of which perform the splicing reaction autocatalytically – without the need of any other molecules. The splicing reaction consists of two consecutive transesterification reactions (Figure 11).[96, 97] The first step comprises a nucleophilic attack of the 2'-OH of a highly conserved adenosine located in a bulge within the sequence (Figure 9, 11). The second step of splicing ligates the 5' and 3' exons and the intron

is released in the typical form of the lasso-like structure called lariat.[98, 99] Substrate recognition relies on the IBS-EBS (intron/exon binding sites) interactions and both of the splicing steps are readily reversible.[100] Alternatively, the first step can also be initiated by the nucleophilic attack of water or a hydroxyl ion, which results in a free 5'-exon and a linear intron/3'-exon intermediate.[101] In this case, the reaction becomes irreversible.[96, 101-103]. Therefore, group II introns catalyze RNA and DNA 5'-hydrolysis and additionally, are able to undergo reverse splicing into new genomic locations:[94] The reverse reaction starts with the binding of the intron via the EBS elements (EBS1 and EBS2) in D1 to the IBS elements in a suitable substrate (5'-exon) next to a splice site. The 3'-OH of the intron attacks the splice site and the intron is inserted via the two reversible transesterification steps. The first 5'-nucleotide of the 3'-exon makes contact to one nucleotide in the subdomain d of D1, thereby describing a third IBS3-EBS3 interaction.[104] Yet, IBS1 is completely sufficient for minimal substrate recognition and the IBS3 (and IBS2) sequence element is only responsible for increased reactivity and for reduction of cleavage at cryptic sites.[105, 106] These simple sequence requirements for substrate recognition represent the high potential of group II intron ribozymes to catalyze other RNA-dependent reactions.[94]

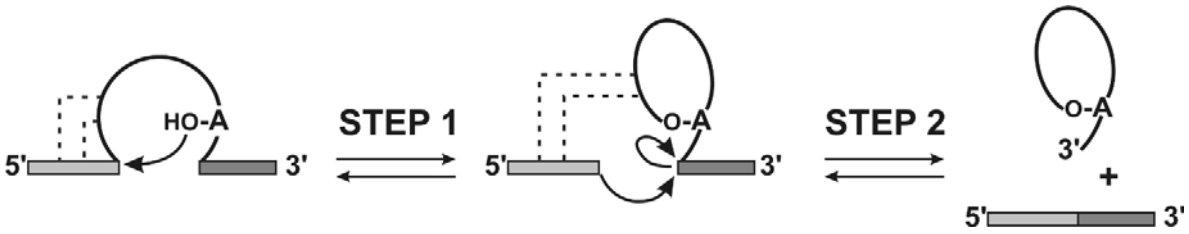


Figure 11. The natural splicing reaction of group II intron ribozymes. The reaction comprises two sequential transesterification steps. The first step consists of a nucleophilic attack by the 2'-OH of the conserved internal branch-point adenosine on the scissile phosphate at the 5'-splice site. In the fast second step, the newly generated 3'-OH of the 5'-splice site attacks the phosphate bond at the 3'-splice site, releasing the intron in the typical form of a lariat and resulting in the ligation of the exons. The exons are shown as bars in different shades of gray, the intron is depicted as a solid black line.

The splicing mechanism shows close similarity to the one performed by the spliceosome in eucaryotes.[107, 108] The striking difference between the spliceosome and group II introns lies in the fact that the catalytic components responsible for splicing reside within the group II intron itself, whereas the spliceosome uses a much more complex set of RNAs combined with proteins for its function. This is a possible example how nature could have developed a specific reaction to a general mechanism during evolution and it would identify group II intron ribozymes as ancient precursors for the spliceosome.

1.7.2 Secondary structure of group II intron ribozymes

Although there is very little sequence conservation among group II intron ribozymes, their secondary structure is divided in a conserved subset of six stem-loop structures or domains (D1-D6) that contain well defined secondary structure elements (Figure 12).[94, 109] All domains are independent folding units[110] and as such can be transcribed as individual molecules. When added *in trans* under the correct conditions, they reassemble into the active ribozyme fold.[111, 112] These interesting abilities make group II intron ribozymes suitable targets to study structure-function and/or metal ion binding-function relationships. The large RNA can be dissected into smaller subunits and several studies on individual domains can be conducted. This information is used in an attempt to understand the greater context of the large RNA, which is hard to study on the whole because of its enormous size. Every domain has a distinct function in folding, catalysis or conformational rearrangement and is involved in intra- and/or interdomain long-range tertiary contacts in the global context of the molecule:

- Domain 1 (D1) is the largest of all intronic domains and serves as the molecular scaffold for the docking of the other domains. D1 recognizes the 5'-exon through two EBS-IBS (exon/intron binding site) base pairing interactions.[109, 113] and the 3'-exon via a one-nucleotide contact EBS3/IBS3.[104] D1 forms intradomain tertiary contacts (α-α', β- β') as well as several interdomain tertiary interactions to D2 (θ-θ') and D5 (κ-κ', ζ- ζ').
- D2 and D3 are not strictly required for catalysis[114] but their presence results in enhanced catalytic activity.[109] D2 lacks phylogenetic conservation but it may have a role in stabilization of the overall structure of the ribozyme through tertiary interactions with D1 (θ-θ') and D6 (η-η').[115, 116] D3 is a relatively short stem-loop acting as an important catalytic effector. It binds with relatively high affinity to the rest of the intron and is thought to be involved in extensive tertiary contact network. Nevertheless, only one tertiary contact of D3 to D5 has been characterized so far (μ-μ').[117]
- D4 is the most variable domain of the intron and can even be removed as long as a short flexible stem-loop linker remains between D3 and D5. D4 often contains an ORF for a multifunctional protein (maturase) important for intron mobility and as such results in a domain of enormous size.[118] It also contains the binding site for the maturase[119] as well as other intron-specific splicing factors.[120] In the folded

intron molecule, D4 is located at the surface and therefore is prone to serve as a general protein-binding element.

- D5 is the phylogenetically most conserved domain[95] and the catalytic center of the intron.[94, 109] This stem-loop structure is the best studied domain of the intron and can be parted into two sides, a binding face with atoms involved in tertiary contacts to D1 (κ-κ', ζ- ζ') and D3 (μ-μ') and a chemical face with functionalities important for catalysis.[121] Information about tertiary structure of D5 and location of possible metal-ion binding sites is available. [49, 122]
- D6 contains the conserved bulged adenosine, which is the nucleophile in the first step of splicing and serves as the branch point.[44, 94] The docking site for the branch point is the asymmetric loop in D1 (the *coordination loop*: nucleotides 222-229 and

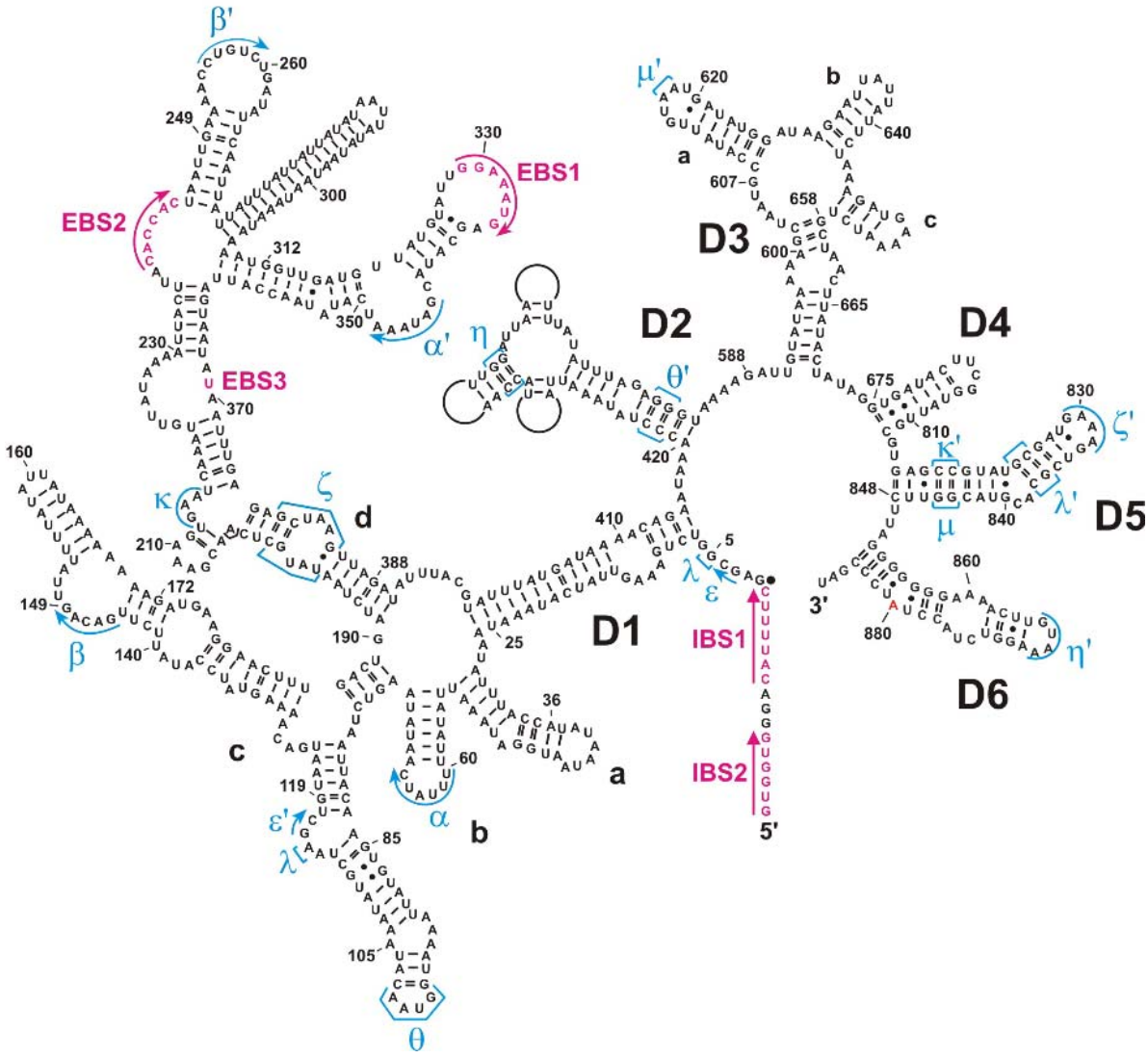


Figure 12. Secondary structure of the group II intron ribozyme *Sc*.ai5γ from the yeast mitochondrial *coxI* gene. Six domains (D1-D6) radiate from a central wheel. D1 and D3 are further divided into subdomains a-d (D1) and a-c (D3). The exon binding sites (EBS1, EBS2, EBS3) and intron bindings sites (IBS1, IBS2) are marked in magenta, the branch-point A in D6 is shown in red, and the tertiary contacts are indicated in blue and by greek letters. The splice site is indicated with a ●.

368-370).[123] D6 forms a tertiary interaction to D2 (η-η') and an NMR solution structure and information about Mg^{2+} binding is available.[124]

1.7.3 Folding of the group II intron ribozyme *Sc*.ai5γ

The group II intron ribozyme *Sc*.ai5γ (yeast mitochondrial intron residing in the *cox1* gene) is able to undergo autocatalytic splicing in the absence of proteins *in vitro* and serves as a model system to study group II intron ribozymes.[97] It displays an optimal catalytic activity at 42°C and under high salt conditions (500 mM KCl, 100 mM $MgCl_2$). Since successful catalysis directly reflects correctly folded and active ribozymes, folding studies are generally performed under these conditions. In the construct *Sc*.D135 that is typically used in folding studies, D2 and D4 are shortened to hairpins and D6 is absent but the original sequence of domains 1, 3 and 5 is conserved (hence the name *Sc*.D135).[125] This ribozyme contains all domains necessary for catalysis and is able to undergo multiple hydrolytic cleavage reactions of a short substrate that contains the IBS1 and IBS2 sequences including the cleavage site of the 5'-exon[126-128], i.e. the substrate 17/7 (*vide infra*).

At first, group II introns were suspected to form an open, solvent-accessible tertiary architecture due to the extended secondary structure and the low sequence conservation among them. Yet, early hydroxyl radical footprinting studies revealed that a striking amount of 40% of the molecule becomes protected from the surrounding solvent, suggesting a compact tertiary structure of group II intron ribozymes.[129] The internalized regions also match the most phylogenetically conserved and mutationally sensitive regions of the intron.

The tight tertiary fold was directly confirmed by equilibrium ultracentrifugation analysis and DMS chemical probing.[125, 130] Furthermore, these studies identified two types of collapsed states. First, a partially collapsed state with a structurally defined secondary structure that forms already in the presence of K^+ ions only. This state is termed the unfolded state. Second, a single, highly compact state of unique tertiary structure that is generated upon Mg^{2+} addition. This fully collapsed species is termed the folded state. Using these definitions, it is possible to exclusively study tertiary folding transitions of group II intron ribozymes.

The local and global features of the folded state were studied under equilibrium conditions to determine thermodynamic stabilities. Local folding was followed by hydroxyl radical footprinting and showed that all regions of the molecules were internalized with a comparable Mg^{2+} dependency at a K_{Mg} of ~20 mM.[125] Furthermore, ultracentrifugation analysis revealed a similar Mg^{2+} dependency also of global molecular compaction (K_{Mg} = 15 mM).[130] A simple and cooperative folding transition was deduced, in which addition of

Mg^{2+} causes a specific and cooperative collapse of the molecule to a single tertiary structure. There is no evidence for any stable intermediate or kinetic trap.[130] Taken together, these findings gave rise to an apparent two-state folding pathway that hinges on the formation and the consolidation of a weak Mg^{2+} binding site.[131]

The folding pathway was also directly monitored using time-resolved biochemical methods that track RNA structure formation through time.[125, 130] Again, local structures would form simultaneously, and both the rate constants for evolution of each footprint and global molecular compaction upon Mg^{2+} addition were similar. Furthermore, the rate constant of acquisition of catalytic activity correlates with the one for tertiary folding (k_{obs} ~1 min^{-1}), proving that the molecule formed upon Mg^{2+} dependent tertiary collapse represents the active species.[132]

Nevertheless, it is unlikely that a molecule of such a large size would fold completely without intermediates. Indeed, deletion analysis revealed that D1 folds and forms with similar rate constant and Mg^{2+} dependencies as the *Sc*.D135 ribozyme.[132] This large domain represents an obligate intermediate along the folding pathway that forms early and is rate-limiting. Domains 3 and 5 rapidly dock to the preformed binding sites, leading to an overall apparent two-state folding of the *Sc*.D135 ribozyme that is dictated by D1 folding (Figure 13).

Just recently, it was found that folding is controlled by a tiny substructure that lies at the center of D1, the κ-ζ element.[56] This conserved, structurally complex junction motif binds and orients D5 in later steps and forms the intron active site. Additionally, the κ-ζ element represents a metal ion binding pocket and is thought to act as a Mg^{2+} dependent switch that initiates the cascade of *Sc*.D135 folding. κ-ζ is perfectly positioned for such a task, since is separates D1 into two halves and is prone to influence the spatial arrangement of long-range tertiary interactions (α-α', β- β') that need to be formed for complete collapse of D1.

The biological relevance of these findings to group II intron function *in vivo* was still uncertain, since all experiments described above were conducted under non-physiological conditions (42°C, 500 mM KCl, 100 mM $MgCl_2$). Therefore, studies under near physiological conditions (30°C, no KCl, 1 mM $MgCl_2$) were performed and revealed a similar folding pathway, albeit with important differences. D1 was found to collapse to a near native intermediate, to which D3 and D5 are capable of docking.[133] Yet, the early collapse of D1 is very slow and D3/D5 docking is unstable. *In vitro*, these problems can be overcome by addition of Mg^{2+}. *In vivo*, this role is possibly taken over by protein cofactors, which have been implicated in numerous aspects of group II intron splicing.[113, 118]

In summary, the rate-limiting step of *Sc*.D135 folding includes formation of an unstable D1 intermediate that is accompanied by the binding of Mg^{2+} to the κ-ζ element.[56] This carefully orchestrated pathway ensures the correct formation of the intron core for subsequent successful splicing at the earliest stages of the folding pathway. Upon formation of the intermediate, the remaining domains fold rapidly into their native positions on the D1 scaffold in an apparently concerted, but independent docking fashion.[132]

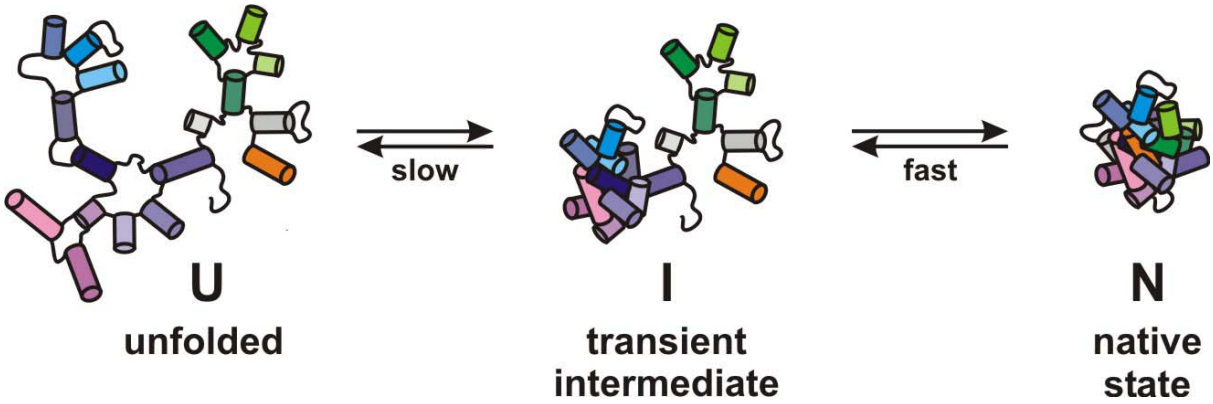


Figure 13. Folding pathway of the *Sc*.D135 group II intron. From the unfolded state U (containing secondary structure elements) a transient intermediate I containing a collapsed D1 is formed in a first slow step. As described in literature so far, a subsequent fast step directly generates the native state N. Helices in D1 are shown in blue and purple cylinders, the helices in D3 are shown as cylinders in shades of green and D5 is shown as an orange cylinder. The truncated helices of D2 and D4 are shown as gray cylinders. Figure is partially adapted from [84].

1.8 Detection of conformational changes

As a RNA molecule folds from the random coil to its native tertiary structure, it experiences extensive conformational changes. Upon metal addition, and therefore charge compensation, portions of an RNA are internalized and no longer accessible from the surface. These local changes can be followed by the hydroxyl radical footprinting method, and also NMR titration experiments report on changes in the local environment. Native gel analysis and sedimentation velocity experiments report on the overall compaction. A very useful technique to directly follow conformational changes as well as the dynamics of specific regions in the molecule relative to each other is the use of fluorescent molecules and their quenching behavior. Therefore - and because of the optically appealing side effect of working with nicely colored solutions - fluorescence is the method of choice used in this study.

1.8.1 Fluorescence Theory[134]

Luminescence describes a family of processes, in which appropriate molecules emit light from exited states created by either physical, chemical or mechanical mechanisms. If the exited states are generated by absorption of ultraviolet or visible light, the phenomenon is termed photoluminescence. Luminescence is formally further divided into fluorescence and phosphorescence, depending on the electronic configuration of the excited state and the emission pathway:

The fluorescence process starts with excitation of a susceptible molecule in the ground state S_0 to an excited singlet state S_1 or S_2 by absorption of an incoming photon within femtoseconds (10^{-15} s). The excited state S1 – and all states in general - contains a number of vibrational energy levels, the lowest of which is usually the most populated and is reached by the molecule via vibrational relaxation (internal conversion) within picoseconds (10^{-12} s). The molecule then returns to the ground state S0 by emission of a photon, a process which occurs on the nanosecond timescale (10^{-9}-10^{-7} s). Due to the loss of energy by internal conversion, the wavelength of the emitted photon is generally shifted to longer wavelengths compared to the incident photon. This wavelength difference is termed the Stokes shift. Usually, the energy levels involved in absorption and emission of light are represented in Jablonski energy diagrams (Figure 14).

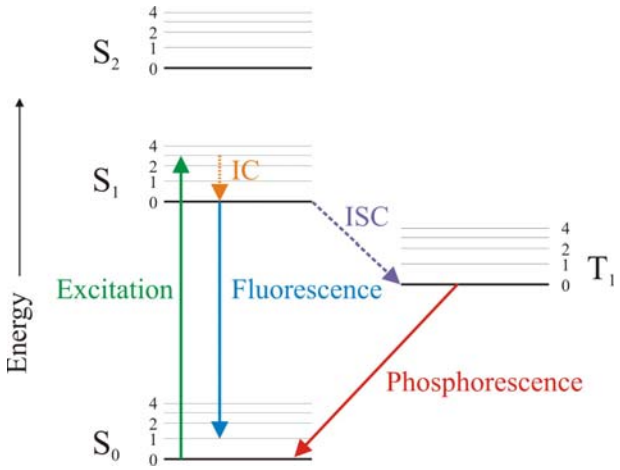


Figure 14. Jablonski energy diagram. The ground state (S_0) and the excited states (singlet: S_1, S_2, triplet: T_1) with vibrational levels are shown. Excitation and emission events are indicated. IC, internal convertion; ISC, intersystem crossing.

Several other relaxation processes with various probabilities compete with the fluorescence emission pathway. The excited state can lose its energy non-radiatively by generating heat or transfer energy to other molecules during collisions. Since excited states are activated states, they are able to undergo chemical reactions with the surrounding and form new species, which are no longer light-emissive and result in loss of fluorescence (photobleaching). In a further event known as intersystem crossing (ISC), molecules convert to the lowest excited triplet states and emit a photon in a process that is now termed phosphorescence. Phosphorescence is less common since it involves spin conversion to produce unpaired electrons, which is an unfavorable (forbidden) process. This is reflected in

slower rate constants of phosphorescence of typically 0.01-10^3 s^{-1} compared to fluorescence (10^7-10^9 s^{-1}). Transitions back to S1 also occur and yield delayed fluorescence.

Molecules able to undergo electronic transitions that result in fluorescence are termed fluorescent probes, fluorochromes or dyes. If these molecules are conjugated to larger macromolecules through adsorption or covalent bonds they are termed fluorophores.[134]

1.8.2 Local structural changes detected by 2-aminopurine (2AP)

2-Aminopurine (2AP) is an example of a fluorescent nucleotide (Figure 15). 2AP is a structural isomer of adenine (6-aminopurine, Figure 2) that can be readily incorporated into an RNA sequence. 2AP is generally a non-perturbing substitution in DNA and RNA because of the structural similarity to adenine.[135] 2AP will form thermodynamically equivalent base pairs with thymine in DNA and uracil in RNA helices and is also capable of forming wobble base pairs with cytosine. In addition, 2AP is highly fluorescent and can be selectively excited in the presence of natural bases.[136] The quantum yield (ϕ, number of times a defined event occurs per photon absorbed) of 2AP fluorescence ϕ_f is sensitive to its microenvironment. The most important factor leading to fluorescent changes were found to be stacking interactions and collisional quenching with neighboring bases.[137] In contrast, the 2AP fluorescence quantum yield is insensitive to base pairing and other H-bonding interactions. Together, these properties make 2AP a suitable fluorescent nucleotide analogue to study subtle local structural changes in nucleic acids. In this study 2AP is used as a reporter group to detect the local structural impact caused by a metal ion that binds to the bulge of an isolated D5 hairpin.

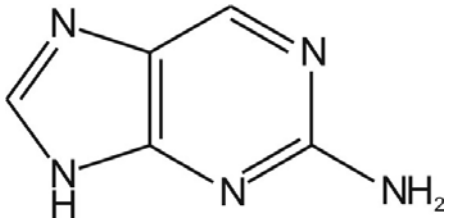


Figure 15. Chemical structure of the fluorescent nucleotide 2-aminopurine (2-AP). λ_{ex} = 310 nm, λ_{em} = 370 nm, ε = 5400 $cm^{-1}M^{-1}$, ϕ_f = 0.68 in H_2O [134]

1.8.3 Förster Energy Resonance Transfer (FRET)

A fluorescence phenomenon of particular interest appears when using a pair of suitable fluorescent probes: The non-radiative energy transfer from a donor to an acceptor molecule by interaction of long-range dipole-dipole couplings. The mechanism is termed Förster or Fluorescence Resonance Energy Transfer (FRET) and is highly dependent on three factors (Figure 16):

1) distance between donor and acceptor molecule (Figure 16a)

2) spectral overlap of donor emission and acceptor excitation (Figure 16b)

3) relative orientation of donor and acceptor dipole moments (Figure 16a)

The theoretical basis of FRET efficiency *(E)* is described by the quantum yield of energy transfer transition E_{FRET}, i.e. the fraction of energy transfer event occurring per donor excitation event:

$$E_{FRET} = \frac{k_{ET}}{k_{ET} + k_f + \sum k_i} \quad (1)$$

k_{ET} is the rate constant of energy transfer, k_f the decay rate constant and k_i the rate constant of all other events that result in non-emissive loss of excitation.

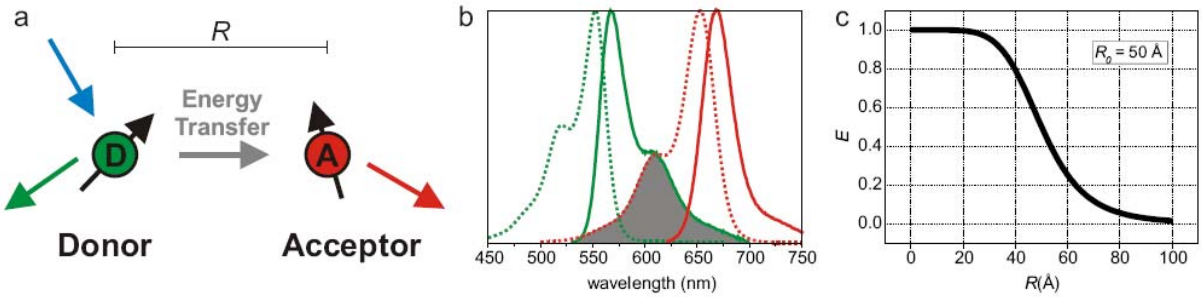


Figure 16. Principle of Förster resonance energy transfer (FRET). **a** Two fluorophores (D, donor; A, acceptor) separated by a distance *R* can generate energy transfer upon excitation (blue arrow) of D. Emission of D is shown as a green arrow, emission of A (generated by energy transfer) as a red arrow, and the black arrows indicate the dipole orientation of the fluorophores. **b** Excitation (dotted lines) and emission (solid lines) of D (green) and A (red). The spectra are normalized for comparison. The spectral overlap of donor emission and acceptor excitation is shown as gray area under the curves. **c** Dependency of energy transfer efficiency *E* on distance *R* at R_0 = 50 Å.

Another description of E_{FRET} is based on the distance between donor and acceptor molecule *(R)*:

$$E_{FRET} = \frac{1}{1 + \left[\frac{R}{R_0}\right]^6} \quad (2)$$

R_0 corresponds to the distance in Å at with 50% of energy is transferred from donor to acceptor and it is termed Förster distance. This number is specific for each pair of dyes and can be calculated by the following equation:

$$R_0^6 = 8.8 \times 10^{-28} \kappa^2 n^{-4} Q_0 J \quad (3)$$

κ is a factor describing the orientation of the dipoles, n the refracting index of the medium, Q_0 the quantum yield of donor in the absence of acceptor and J the spectral overlap integral calculated as:

$$J = \int f_D(\lambda) \varepsilon_a(\lambda) \lambda^4 d\lambda \quad (4)$$

where f_D is the normalized donor emission spectrum, ε_a depicts the acceptor molar extinction coefficient.

All variables in these equations except κ and R are constants and have been determined for numerous fluorescent probes. The dipole orientation factor κ can theoretically vary between 0-4 and is a measurement for the rotational freedom of the fluorescent dye. If the dyes are freely rotating and isotropically oriented in the excited state a constant value of $\kappa^2 = 2/3$ can be assumed. This leads to the fact that the only variable parameter is *R*, the distance between donor and acceptor at any time. Therefore, changes in FRET report on the relative distance changes between a donor and an acceptor molecule and the technique can be used as a spectroscopic ruler in the range of typically 2-8 nm distance between the donor and acceptor molecule (Figure 16c).[138, 139]

A common dye pair used in FRET experiments – especially for single molecule studies (*vide infra*) on RNA[140, 141] – is the combination of the cyanine dyes Cy3 and Cy5 (Figure 17a, b). They are water-soluble, bright and photostable dyes that fluoresce in the green (Cy3, λ_{em} = 565 nm) and red (Cy5, λ_{em} = 665 nm) and can be coupled to nucleic acids or proteins over their aliphatic nitrogen-side chains R.[142]

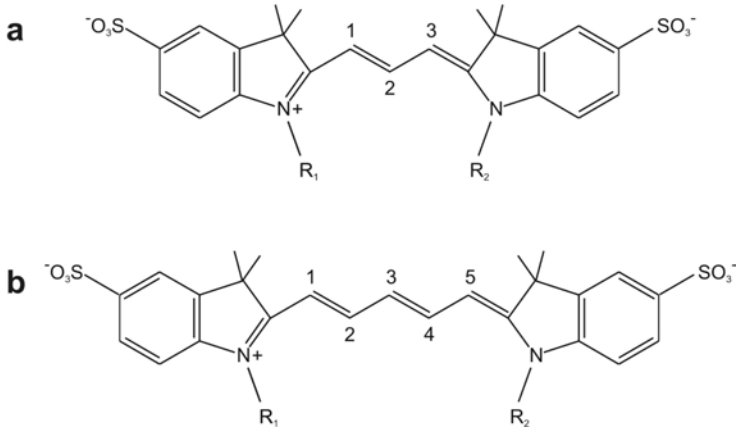


Figure 17. The cyanine dyes. **a** Cy3 (λ_{ex} = 540 nm, λ_{em} = 565 nm, ϕ_f = 0.13) and **b** Cy5 (λ_{ex} = 640 nm, λ_{em} = 665 nm, ϕ_f = 0.18). The linker length is indicated, whereby the number of linking atoms determines the cyanine dye name.

1.8.4 Global conformational changes at the single molecule level

The body of this work is composed of detection of conformational changes by FRET. The range, in which FRET can be used as a powerful spectroscopic ruler (Figure 16c), matches the global dimension of the ribozyme under study. By positioning of dyes at various sites on the molecule it is possible to follow the relative changes of individual regions of the molecule during the folding process.[64, 143]

Bulk measurement of many molecules at the same time result in elevated and easily detectable levels of fluorescence signals and report on the average evolvement of FRET during folding. Apart from the wealth of information that can be gained by these experiments,

one major drawback is the fact that the existence of short-lived intermediate structures will not be detected, since their contribution is only weak and as such will be overpowered and can not be detected. A very powerful method of following individual molecules in real-time is achieved by single molecule spectroscopy.[144-147] This technique has more recently also been applied to RNA.[141] By looking at an individual molecule it is possible to access detailed dynamic information, i.e. detect stochastic fluctuations between distinct conformational states and measure the corresponding rate constants in real time. Additionally, potential transient intermediate and parallel folding pathways can be resolved and identified independently.[148]

1.9 Thesis Outline

This work deals with different aspects of metal ion binding to an RNA structure. The RNAs under investigations are mainly derived from the group II intron ribozyme *Sc*.ai5γ from *Saccharomyces cerevisiae* (Baker's yeast), but also from a related thermophile ribozyme derived from *Av*.groEL from *Azotobacter vinelandii*. Chapter 2 describes an assay to generally detect the metal ion binding sites in the group II intron ribozyme *Av*.groEL. In Chapter 3, the impact of binding of one metal ion on the local structure of isolated D5 is studied by monitoring 2AP-fluorescence. Chapter 4, which represents the main body of this work, deals with the global structural changes upon metal ion binding of the ribozyme *Sc*.D135 by measuring FRET at the single molecule level.

2. Metal ion binding sites in the group IIB intron *Av*.D135

2.1 Introduction

2.1.1 The group II intron ribozyme *Av*.groEL

Group II intron ribozymes were originally described in genomes of mitochondria and chloroplasts, where they interrupt any type of essential genes.[149] The existence of this group of introns in bacteria was discovered only a decade later.[150] But since then, the bacterial group II intron ribozymes have been found to be widespread and somewhat more diverse than their counterparts in organelles, consistent with the idea of a bacterial origin of group II introns.[151] Bacterial group II intron ribozymes are found at many different locations except within functional housekeeping genes and are often associated with mobile genetic elements.[151-153]

In general, group II introns are classified into subgroups based on comparative analysis of secondary and tertiary structural elements.[154] The self-splicing mitochondrial introns from *Saccharomyces cerevisiae* (*Sc*.ai5γ) and *Pylaiella littoralis* (*Pl*.LSU/1 and 2) belong to subgroup IIB1. Other members of the group IIB1 are the uncharacterized bacterial introns from *Escherichia coli* (*Ec*.p0157), *Pseudomonas sp.* and *Xylella fastidiosa*.[155]

The intron *Av*.groEL (also known as *Av*.hsp60) from the nitrogen-fixing bacterium *Azotobacter vinelandii* was identified in 1993[150] and turns out to be closely related to the subgroup IIB1 yet containing unusual features that diverge from typical subgroup IIB1 signatures:

- Although it is a bacterial ribozyme, *Av*.groEL interrupts the termination codon of the essential gene groEL, which encodes a heat shock protein (hsp60 chaperone).[155]
- *Av*.groEL displays a canonical group IIB secondary structure of 6 domains but is extended with a 292 nt insertion close to the 5' end of the intron (Figure 18). The intron contains a large ORF of 1692 nts that encodes a polypeptide with reverse transcriptase, maturase and endonuclease modules. *Av*.groEL shows a G-C basepair content of 59.2%, which is significantly higher than what is found for other self-splicing group II introns.

- Group II intron ribozymes generally require non-physiological conditions for *in vitro* activity, i.e. high salt concentrations and temperatures.[96, 156, 157] The construct *Av.*groEL without the ORF shows an optimal activity at particularly high temperatures of 52-55 °C,[155, 158] and a more streamlined version (no ORF, no 292 nt insert) even prefers 65°C.[155, 158] Hence, *Av.*groEL displays an unusually high thermostability and a requirement for thermal activation of splicing.

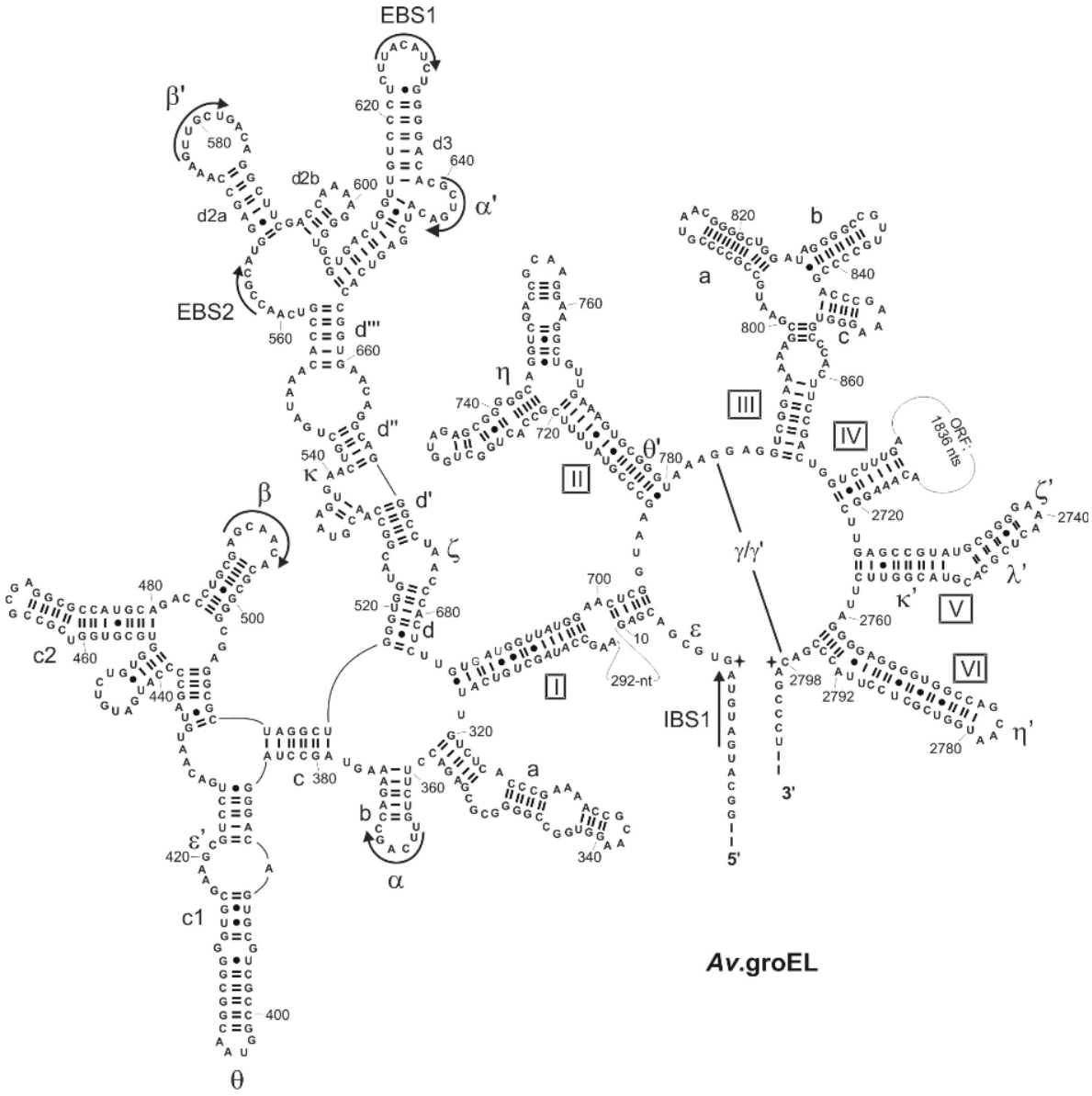


Figure 18. *Av.*groEL intron sequence and secondary structure map. Domain are labeled with boxed roman numbers (I-VI) and tertiary interactions are indicated in greek letters. Arrowheads indicate tertiary contacts formed by Watson-Crick type basepairing. The 292 nt insertion in D1 and a part of D4 (1836 nt), which includes the ORF, are represented as loops.

Taken together, *Av.*groEL represents an interesting exception to bacterial group II introns that seems to be optimized for high temperatures and is potentially involved in heat shock regulation. The sequence of *Av.*groEL contains an high G-C content (59.2 %)[159] compared to *Sc.*ai5γ (20.8 %) or *Pl.*LSU1 (40.9 %).[160] However, the reason for this does not simply

rely in the G-C content as the correlation between elevated G-C content and increased thermal stability of RNA (due to three hydrogen bonds in a G-C basepair compared to only two in A-U basepairs) is not straightforward. There is no single mechanism that explains the effects of G-C content on the function of large RNAs. Instead, increased stability was described to be caused by more homogeneous folding to the native state[161] or an increase in secondary structure content.[162] Further factors like coordination of water molecules or – interesting and important in respect of this work - metal ion-RNA interactions may play a critical role, too.[162, 163]

2.1.2 Principle of terbium cleavage

Location and function of metal ions in folded RNA molecules have been studied using a variety of techniques. Historically, hydrolytic metal ions like Pb^{2+}[164, 165] or oxidative ions such as Fe^{2+}[166] have been applied to probe divalent metal ion binding sites. Although these approaches have led to valuable preliminary results, their significance is limited by the altered geometric or coordination preferences of these metal ions compared to Mg^{2+} or by the requirement of the experiments to be conducted under nonphysiological conditions. In further studies, phosphorothioate rescue experiments were conducted to test for direct coordination between metal ions and pro-Rp phosphoryl oxygens in the RNA using metal ions like Mn^{2+} or Cd^{2+}.[33] The drawback here lies in the fact that this type of coordination does not always appear to be required in folding of group II intron constructs and therefore, positions of metal ion binding detected with phosphorothioate rescue experiments might not generally be involved in metal ion mediated interactions.[167]

Lanthanide(III) ions have been widely used to map metal ion binding sites in RNA structures.[165, 168, 169] In particular, Tb^{3+}-mediated footprinting has proven to be a versatile technique to probe ribozyme secondary and tertiary structure and metal ion binding (*vide infra*).[170-172] Hexacoordinate Tb^{3+} has a similar ionic radius of 0.92 Å compared to Mg^{2+} (0.72 Å)[173] and shows the same preferences for coordination of oxygen ligands.[174] Therefore, Tb^{3+} is potentially able to bind to similar sites on a polyanionic RNA as Mg^{2+}. Tb^{3+} shows a binding affinity that is two to four orders of magnitude higher[170, 171, 175] and thus, Tb^{3+} can easily replace Mg^{2+} in its specific binding sites already in low concentrations (Figure 19). Once bound to the RNA, the transiently deprotonated aquo complex $Tb(OH)(aquo)^{2+}$ abstracts a proton at the 2'-OH group from a nearby nucleotide.[165, 176] The resulting 2'-oxyanion attacks the juxtaposed phosphodiester in a nucleophilic fashion, leading to slow backbone cleavage. This reaction takes place at physiological pH, since the

pK_a of Tb^{3+} is lowered to 7.9 compared to Mg^{2+}, for which pK_a = 11.4.[177] Efficient local backbone scission occurs at every Tb^{3+} binding site on the RNA with correct geometry (Tb^{3+} binding usually in the minor groove)[172] and results in a distinct set of cleavage products of different length.[172, 175]

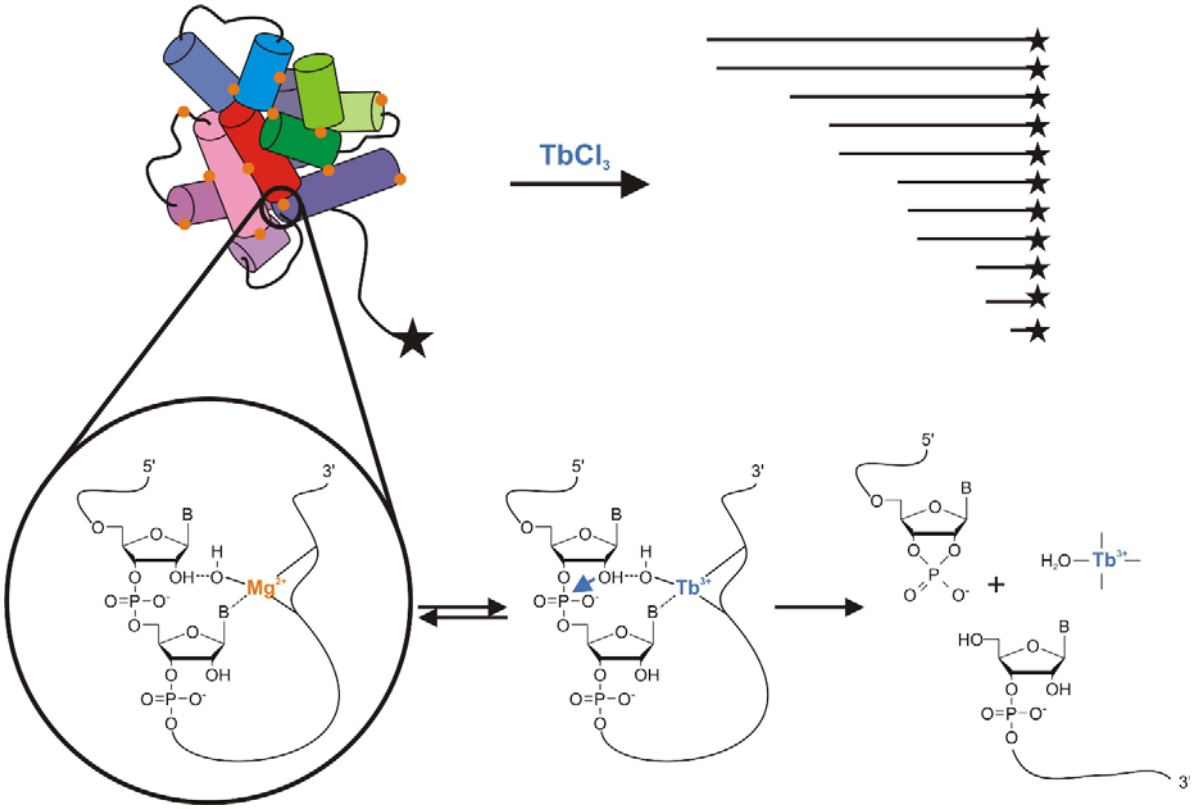


Figure 19. Principle of Tb^{3+} cleavage. Mg^{2+} bound in its binding sites on an RNA is replaced by Tb^{3+}, which results in cleavage of the RNA backbone at the site of metal ion binding and gives rise to a distinct set of cleavage products.

At low (micromolar) concentrations, Tb^{3+} binds preferentially to RNA surfaces of high local charge density, thereby revealing potential Mg^{2+} binding sites. At increased (millimolar) concentrations, Tb^{3+} cuts the phosphodiester backbone in a largely sequence-independent manner, preferentially in single-strand or non-Watson-Crick basepaired regions. Therefore, Tb^{3+} mediated RNA cleavage at low concentrations is used to map specific metal ion binding sites, whereas at higher concentrations it serves to footprint secondary and tertiary structure.[170]

Interactions of lanthanide(III) ions and RNA can be studied in further assays that prove the binding of lanthanides to RNA and show the versatile potential of these remarkable metal ions. Replacement of Mg^{2+} with Tb^{3+} has been found to result in inhibition of catalysis of the hairpin and the hammerhead ribozymes.[171, 178] Furthermore, Tb^{3+} exhibits distinct luminescence properties when bound to RNA. Tb^{3+} shows a small extinction coefficient of direct absorption, but when proximally bound to nucleobases Tb^{3+} can emit sensitized

luminescence. This occurs when Tb^{3+} is bound to RNA and nearby guanines transfer energy to the lanthanide ion upon excitation.[179] Using this property, Tb^{3+} is applied to conveniently measure metal ion binding affinities.[171, 178, 180, 181]

2.1.3 Aim of work

Major metal ion binding sites in the group II intron core have been identified and described for *Sc*.ai5γ.[175] Here, analogous Tb^{3+} titration were performed on *Av*.groEL. Data from these two related but distinctly different ribozymes were compared to reveal shared as well as individual metal ion binding properties.

Previous studies in our lab independent of this work failed to identify cleavage sites in the 5' half of the molecule. This work describes a successful test run to find metal ion binding sites in the 3' half of the large RNA molecule. Since data for the 5' half of the molecule is still missing, the data set is not complete yet.

2.2 Results

2.2.1 The RNA construct *Av*.D135

The sequence of the RNA construct used in these metal ion mapping experiments starts at the first nucleotide of the *Av*.groEL intron and is extended with an unrelated 37 nt tail at its 3'-end. The 292 nt insertion close to the 3'-end of the ribozyme in D1 was removed. The original sequence of D3 and D5 remained preserved, whereas D2 and D4 were replaced by short hairpins that are capped by UUCG tetraloops. The resulting construct was called *Av*.D135. The construct is analogous to *Sc*.D135 from *Sc*.ai5γ (Figure 12), which is known to subscribe all the necessary components of the active site of the ribozyme for splicing.[128]

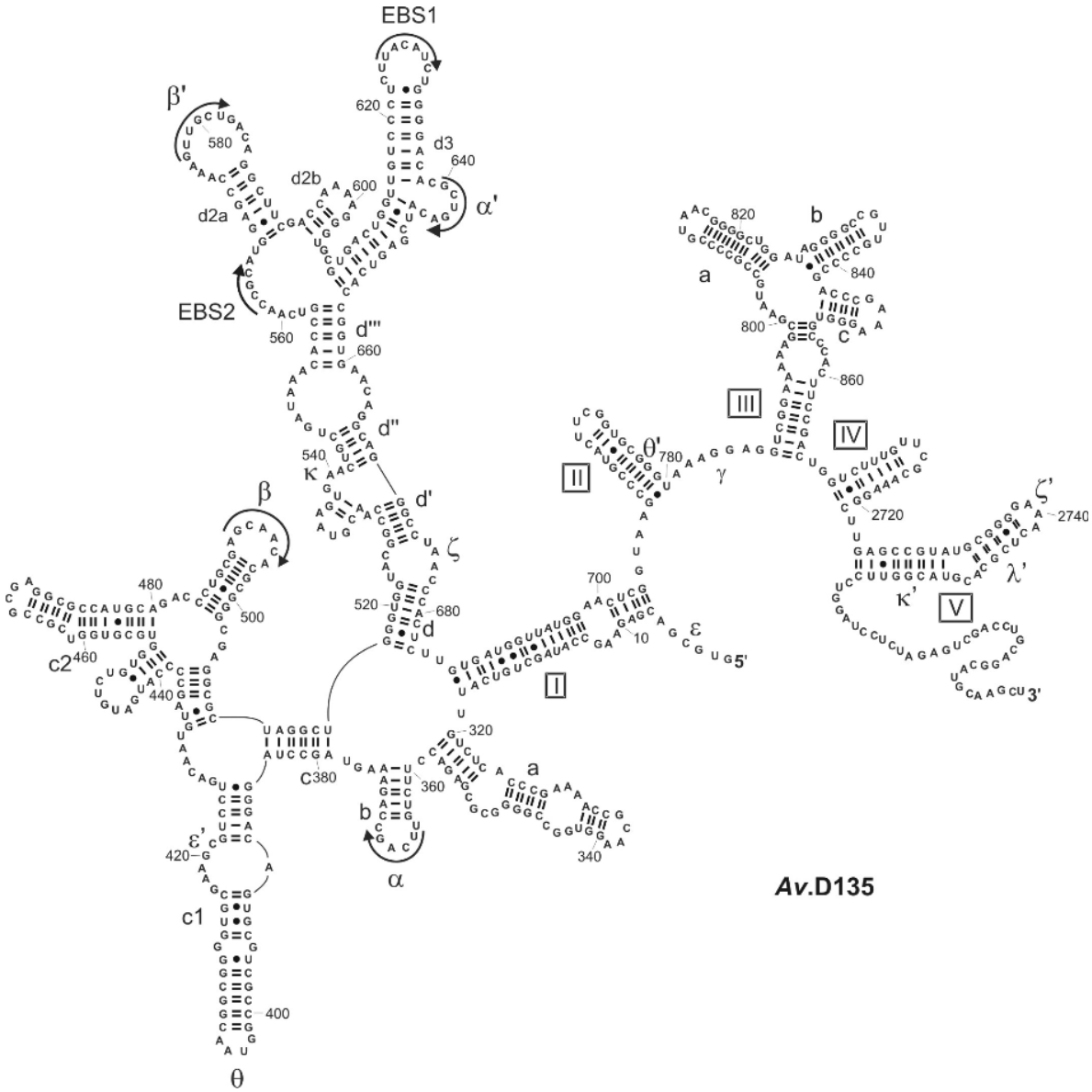


Figure 20. Sequence and secondary structure map of construct *Av*.D135. Domains are labeled with boxed roman numbers (I-V) and tertiary interactions are indicated using greek letters. Arrows indicate tertiary contacts formed by Watson-Crick type basepairing.

2.2.2 Terbium(III) cleavage

For detection of cleavage bands resulting from strand breaks at sites of Tb^{3+} coordination, *Av*.D135 was 3'-end labeled with α-^{32}P-dCTP. The labeling products were purified via denaturing polyacrylamide gel electrophoresis (PAGE) and the major product band was isolated and separated from side bands as effectively as possible. Side products originate from the transcription, whereby T7 also generates products that are shorter by some nucleotides, and from an oligo splint used in the labeling reaction that allowed incorporation of either one or two labeled radionucleotides due to an unfortunate error in the nucleotide sequence of the oligo splint. Labeled *Av*.D135 was then prefolded in $MgCl_2$ before Tb^{3+} was added and the cleavage reactions were allowed to advance for 1h on ice. Cleavage products were separated and subsequently imaged on denaturing polyacrylamide gels. Since full separation of long RNAs differing by only 1-2 nucleotides is never completely achieved, the cleavage bands would often turn up as multiple (mostly double) rather than a single band and complicate the observed cleavage pattern. In addition, the G-specific T1 ladder of *Av*.D135 already contains a large number of products due to the elevated amount of G's in the *Av*.D135 sequence, which leads to cleavage patterns that are not easily identified and complicates the assignment. Nevertheless, the assignment was possible for ~200 nts starting from the 3'-end of the ribozyme, whereas it became increasingly difficult in D1 as bands were not resolved as well anymore. In addition, comparison to the cleavage pattern from the related *Sc*.ai5γ sequence successfully helped to assign some more cleavage bands to nucleotide numbers in the 3' part of D1.

2.2.3 Metal ions in the catalytic core

D5 is known to be at the heart of the catalytic core of group II intron ribozymes.[121, 182-185] A construct of D1 and D5, in which D2 and D4 are shortened to hairpins and D3 is removed but interdomain linkers are preserved in original sequence, is a minimal form of an active group II intron ribozyme derived from *Sc*.ai5γ.[132] The tertiary contacts ζ-ζ'[30] and κ-κ'[167] link D5 to the D1 scaffold and a further contact λ-λ' brings the active site elements in close proximity to the 5'-splice site.[185] Therefore, it was of great interest to pin down metal ion binding sites in this apparently central domain within *Av*.D135.

The most prominent metal ion binding site in D5 is found around the bulge nucleotides A2747-G2749 (Figure 21). The conserved appearance of this cleavage site in *Av*.D135 and *Sc*.ai5γ[175] is another strong proof that a specific metal ion binds to the bulge region and is a

component of the active site of the ribozyme molecule. As described previously,[175] metal ion cleavage extends to adjacent base pairs in stem 2 of D5 also in *Av*.D135. These nucleotides are part of the λ-λ' tertiary interaction that connects the chemical face of D5 with the 5' splice site.[185] A metal ion appears to mediate this long range interaction essential for catalysis. The D5 tetraloop is only cleaved at high concentrations of Tb^{3+}. This indicates weak metal ion binding and minimal single strand character at this site, approving the binding of the tetraloop to its receptor site via ζ-ζ'[30] in the folded molecule.

The domains are linked by short single-strand joiners, the catalytic role of which is largely unknown. A prominent metal ion binding site is located in the junction linking domains 2 and 3 (J2/3). This purine-rich joiner is well conserved[154] and has been implicated to be another important element of the active site.[186] G784, which usually is a conserved A in most other group II introns, is involved in formation of the γ-γ' tertiary

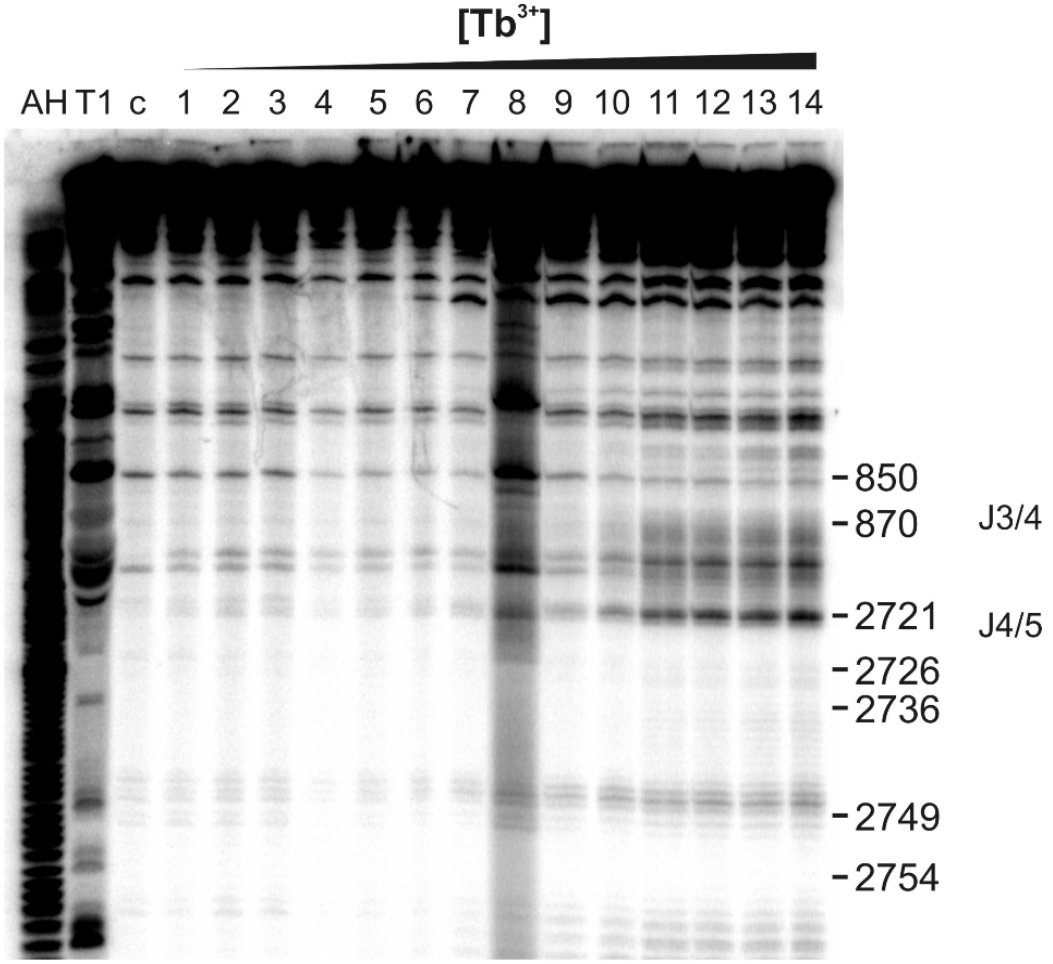


Figure 21. Tb^{3+} cleavage sites in domains 4 and 5 resolved on a 18 % denaturing polyacrylamide gel. Lanes 1-14 represent samples incubated with Tb^{3+} at concentrations of 0,0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.75, 1, 2.5, 5, 7.5, 10 and 25 mM. Lane c shows RNA without any M^{n+} added, lanes AH and T1 represent alkaline hydrolysis and T1 endonulease ladders, respectively. Assignment of cleavage bands to nucleotides and regions in the *Av*.D135 sequence is shown to the right.

contact. Formation of the γ-γ' interaction leads to recognition and cleavage of the 3'-splice site.[187, 188] Divalent ions have previously been described to take part in reaction chemistry at the 3'-splice site[189] and the results presented here support this finding. The adjacent conserved residues G785 and A786 exhibit further weak metal ion binding. These nucleotides have been found to be essential for the second step of splicing.[190] They have also been reported to photo-crosslink with D5, contacting nucleotides U2756 and U2757, which are opposite the catalytic AGC triad in stem 1 of D5,[102] or even the bulge in D5.[191] A metal ion binding site is located at this junction. As mentioned above, J2/3 is involved in catalysis and therefore, our result for this site implies a possible role of a metal ion participating in catalysis.

2.2.4 Metal ions at sites involved in exon recognition

D1, the largest domain of the intron, is absolutely essential for catalysis and provides the scaffold for the assembly of the other domains into the catalytic active ribozyme. It displays an extensive network of inter- and intradomain tertiary interactions. It also harbors the exon binding sites EBS1 and EBS2. D1 is responsible of substrate recognition and places the 5' splice site within the intron core.[192] Metal ions have been described to bind to nucleotides flanking EBS1 and the β-β' interaction using Tb^{3+}-cleavage.[110, 175]

Assignment of Tb^{3+} cleavage sites to nucleotide positions within D1 becomes increasingly difficult for 3'-labeled *Av*.D135 and is impossible for the 5'-half of D1 since degradation products can no longer be nicely separated due to size and overlay (Figure 22). The T1 ladder also becomes crowded and sequence assignment is difficult. Yet comparison of gels from *Av*.D135 and *Sc*.ai5γ[175] reveals similar Tb^{3+} cleavage patterns of the two related ribozymes and allows an estimate for the nucleotide assignment in the 5'-half of D1. Using this approach, nucleotides flanking EBS1 and the β-β' interaction were identified to be metal ion binding sites. Unfortunately, Jd"/d"', which has been described to comprise a further exon binding site EBS3,[104] overlays with a prominent cleavage band and conceals Tb^{3+} cleavage information.

2.2.5 Metal ions involved in further tertiary contacts

Another important intradomain contact within D1 is described by α-α'. Formation of this interaction places distant secondary elements in close proximity, a feature that forms upon

excessive compaction of the RNA molecule. Cleavage signals adjacent to α' indicate a metal ion that mediates this long range tertiary contact (Figure 22).

The κ-κ' and ζ-ζ' tertiary contacts have lately become the center of attention for folding of group II intron ribozymes.[56] No metal ion binding effects are detected in Tb^{3+} mediated footprinting, neither in a previous study on *Sc*.ai5γ nor in *Av*.D135 presented here. Recently, the κ-ζ element in D1 has been described to be a metal ion binding pocket and to act as a Mg^{2+} dependent switch for initiation of folding using time resolved nucleotide analog interference mapping (NAIM).[56]. This absence of an important ion binding site reflects limitations in the Tb^{3+} cleavage assays. For successful backbone scission by Tb^{3+}, the 2'-OH group needs to be precisely aligned (Figure 19). This conditions is not met throughout the entire intron that contains excessive secondary and tertiary structure and therefore, conformational constraints (also see Section 2.3).

2.2.6 Metal ions in regions of undetermined importance

A number of cleavage sites occur in D3, which acts as a catalytic effector for the splice reaction (Figure 22). D3 is not specifically required for catalysis but greatly enhances the reaction rate, probably through direct involvement of D3 functional groups to optimize the active site or in an allosteric fashion.[186] Mainly the single-strand regions in D3 are cleaved at millimolar Tb^{3+} concentration, including the most conserved substructure of D3, which is the A-rich internal bulge within the basal stem.[95] The recently found interaction μ-μ'[117] that connects nucleotides of the pentaloop in D3 with the backbone of G2765 in D5, does not seem to involve a binding site for metal ions in this Tb^{3+} cleavage assay. The strongest cleavage site in D3 resides in between stem c and b, the same site that was also prominently cleaved in D3 from *Sc*.ai5γ.[175]

Similarly, weak cleavage occurs all throughout D4 (Figure 21, 22). D4 is very large in the original sequence and harbors the ORF for group II intron related proteins.[118] D4 also serves as a protein binding platform and therefore is found at the surface of the molecule.[119, 120] The size of this domain varies considerably throughout group II introns, but D4 is always found at its specific location and never completely removed.[94] Since not being involved in catalysis, D4 is reduced to a hairpin of unrelated sequence in the construct *Av*.D135. Weak Tb^{3+} cleavage is found throughout the hairpin, being more pronounced in the UUCG tetraloop and the adjacent domain linkers. If metal ion binding to the stem of D4 is

preserved in the hairpin that replaces D4, this might potentially indicate a role of D4 and metal ions in the overall ribozyme architecture.

All linker regions of the central wheel of the secondary structure appear to be involved in metal ion binding (Figure 22). The linkers are usually rather conserved in sequence and known to be involved in catalysis, although their function is not completely understood. Strong binding sites are found in J2/3, as discussed previously, and J4/5.

2.2.7 Comparison of cleavage sites in *Av*.D135 and *Sc*.ai5γ

Comparison of the cleavage patterns from *Av*.D135 and *Sc*.ai5γ[175] reveals a closely related map of metal ion binding sites (Figure 23). Over all, D3 shows higher metal ion binding than D4 in *Sc*.ai5$\tilde{\gamma}$. In contrast, binding of metal ions to *Av*.D135 is more pronounced in D4 than D3 (Figure 22). This observation is possibly an artifact that originates from introduction of the unrelated hairpin sequence in D4 and its orientation in the RNA constructs used in experiments. On the other hand, the cleavage in D3 of *Av*.D135 seems to be reduced and shifted to higher concentrations compared to *Sc*.ai5γ. This is due to the elevated G-C basepair content in *Av*.D135 as helices consisting of A-U basepairs are slightly more efficiently cleaved compared to double stranded regions of G-C basepairs. This effect is especially pronounced in the short stems a, b and c of D3. Nevertheless, the strong metal ion binding site in between stems b and c is clearly preserved in both ribozymes.

There is one prominent absent cleavage site in *Av*.D135 compared to *Sc*.ai5γ. Weak Tb^{3+} cleavage is detected throughout J3/4 of *Av*.D135, whereas there is a strong cleavage site 3' to D3 in the same linker for *Sc*.ai5γ (Figure 23$\tilde{}$) This result indicates an interesting and potential significance of the linker regions for metal ion binding and function or architecture of the intron.

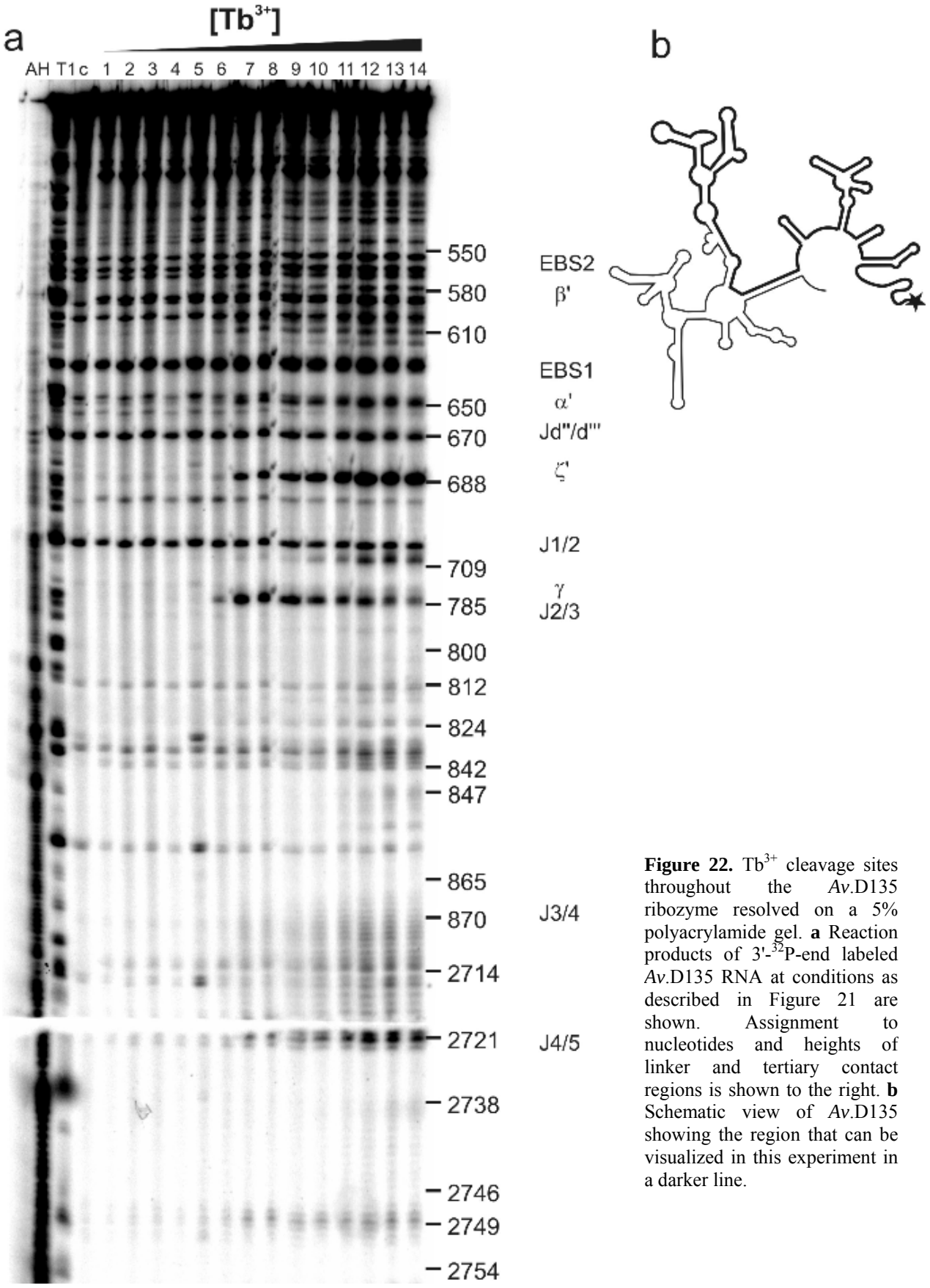


Figure 22. Tb^{3+} cleavage sites throughout the *Av*.D135 ribozyme resolved on a 5% polyacrylamide gel. **a** Reaction products of 3'-^{32}P-end labeled *Av*.D135 RNA at conditions as described in Figure 21 are shown. Assignment to nucleotides and heights of linker and tertiary contact regions is shown to the right. **b** Schematic view of *Av*.D135 showing the region that can be visualized in this experiment in a darker line.

2.3 Discussion and Conclusion

Tb^{3+} cleavage experiments were carried out to locate metal ion binding sites in the unusual ribozyme *Av*.D135. The detected cleavage positions are depicted on the secondary structure of *Av*.D135 in Figure 23. Most single stranded regions were cleaved at high Tb^{3+} concentrations in agreement with the basic concept of Tb^{3+} cleavage.[180] Predominantly, specific metal ion binding sites were found at positions that are part of the catalytic core, which includes D5 and the J2/3 linker. These results are in clear support of the idea that D5 represents a metal ion binding platform.[49, 93, 193]

Due to degradation bands and increasingly reduced spacing of cleavage bands and therefore overlaying signals, assignment of cleavage sites in D1 became difficult. Nevertheless, comparison with results for *Sc*.ai5γ[175] allowed to locate important sites involved in ribozyme reaction chemistry and detect metal ion binding, i.e. at EBS1. Our results link metal ions to sites of exon recognition and are consistent with the implication of metal ions in the chemical mechanism of splicing.[189]. Naturally, ribozymes are always associated with metal ions (Section 1.4) and *Av*.D135, like *Sc*.ai5γ, is also a metalloribozyme if referring to the definition of metalloribozymes that directly connects a divalent metal ion to ribozyme catalysis. [38, 42]

Further cleavage at sites adjacent to nucleotides involved in tertiary contacts were found near α' and β'. Both of these elements form Watson-Crick type basepairs with their counterparts that are distant in primary and secondary structure. Metal ions here might be involved in stabilization of the ribozyme tertiary fold, reflecting the alternative role of metal ions in catalysis and folding of ribozyme structures.

High affinity metal ion binding sites in large RNAs can be identified in lanthanide(III)-mediated hydrolysis experiments. The lanthanides form metal hydroxides that cleave RNA through a mechanism that is highly dependent on local geometry.[164]. Good candidates for cleavage are the backbone of single stranded regions, dynamic loop structures and the widened major groove of non-Watson-Crick basepairs. It is certain that other strong metal ion binding sites remain undetected during the assay. Especially sites within RNA helices, such as the major groove of in G-C rich duplexes or tandem G-U wobble pairs, will be missed. The reduced cleavage in D3 of *Av*.D135 that contains many more G-C basepairs in the stems a, b and c compared to *Sc*.ai5γ is a good example to illustrate this.

For a complete set of metal ion binding sites, analogous experiments on 5'-labeled *Av*.D135 are needed to reveal metal ion binding sites in D1 in much greater detail. But already

with the present preliminary data it becomes apparent that the general cleavage patterns of *Av*.D135 and *Sc*.ai5γ are closely related, indicating a general usage of metal ions in catalysis for and folding of group IIB1 intron ribozymes. Interestingly, the only difference in metal binding observed is the lack of a strong metal ion binding site at the linker region between D3 and D4 in *Av*.D135 when compared to *Sc*.ai5γ, pointing to a possible importance of the junction regions for function of the ribozyme that was so far largely ignored. A finding like this, combined with results from many more experiments using different biochemical techniques, might help to explain the varying features of the large and complex group II introns originating from different hosts in the long run.

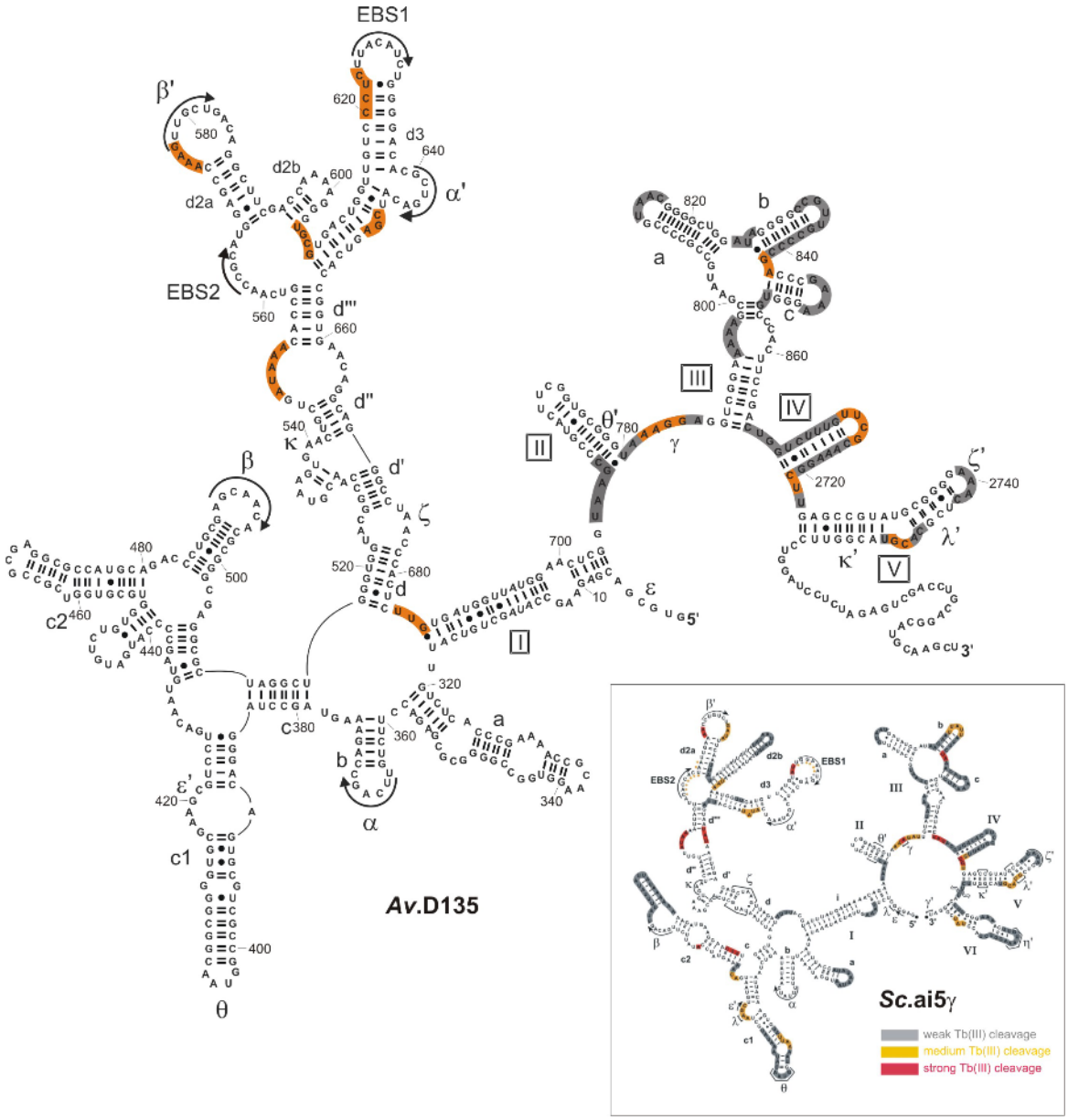


Figure 23. Sites of Tb^{3+} cleavage superimposed on the secondary structural representation of *Av*.D135. The *Av*.D135 construct is derived from *Av*.groEL (Figure 20). Domains 2 and 4 are shortened to hairpins, but the numbering system corresponds to the intact intron. Regions where metal cleavage was observed are shaded in gray (weak cleavage 5'000-13'000 cpm.) and orange (distinct cleavage, 13'000-30'000 cpm). Inlet: Sites of Tb^{3+} cleavage in the *Sc*.ai5γ intron as published in [175] for comparison.

3. Domain 5: Impact of metal ion binding on local structure and stability

3.1 Introduction

Domain 5 (D5) is a key component in the active site of group II intron ribozymes.[109, 127] Together with D1, D5 is the only domain absolutely required for catalysis. The central role of D5 for the splicing reaction is further reflected by the fact that it is the most phylogenetically conserved part of group II introns.[95] D5 generally consists of 34 nucleotides and forms a hairpin, whose stem is separated by an internal dinucleotide bulge into a lower helix 1 and an upper helix 2. The hairpin structure is typically capped by a GNRA tetraloop (Figure 24a).[194, 195]

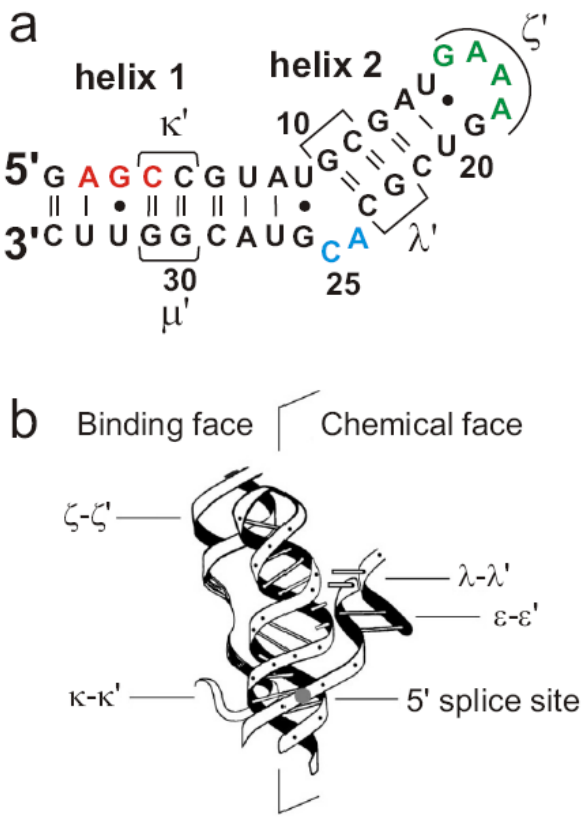


Figure 24. Structural models of D5 from *Sc*.ai5γ. **a** Secondary structure depiction. The catalytic triad is colored in red, the tetraloop in green, and the dinucleotide bulge in blue. **b** Putative structure model of the D5 active site showing the partitioning into the binding and chemical face with the corresponding tertiary contacts as described in the text. The 5' splice site is shown in gray and tertiary contacts are indicated in greek letters. Figure b is adapted from [185].

Many functional groups within D5 are involved in binding to D1 or are important for catalysis.[121, 182, 185] The residues important for binding to D1 project into the minor groove and define the continuous "binding face" of D5. On the opposite side of the molecule, the individual functionalities involved in catalysis span a major groove region and define the "chemical face" (Figure 24b).[184] Two tertiary contacts are formed from the binding face of D5 to D1, i.e. the κ-κ' [167] and the ζ–ζ' [30] tetraloop-receptor interactions. Another tertiary contact from D5 to D1 is found on the chemical side of D5 and involves the λ–λ' interaction, which places the chemical face of D5 close to the 5'-splice site.[185] More recent results show that D5 interacts also with D3, the catalytic effector, via the μ–μ' contact.[117]

The three nucleotides AGC, located close to the 5'-end of D5, are almost invariant throughout group II intron ribozymes. This triad is known to be crucial for catalysis.[183, 184, 196] Therefore, it is frequently called the *catalytic triad*. Another important region for splicing within D5 is the dinculeotide bulge separating the two helical stems.[197] Both the AGC-triad and the dinucleotide bulge as well as the capping tetraloop were described to specifically bind metal ions.[49, 175, 193]

Structural data on D5 from *Sc*.ai5γ is available in form of a solution NMR-structure [49] as well as from X-ray crystallography.[122] Both structures show a stable D5 hairpin with helical A-form regions that are separated by the bulged nucleotides. Considerable differences between the two mentioned structures appear in this dinucleotide bulge. Hydrogen bonding from U9 to G26 and bulged A24 and C25 are proposed from the crystallographic data (Figure 25a). In contrast, the NMR structure shows a new RNA folding motif, in which the U9, A24, C25 and G26 nucleotides are unpaired. In the latter model, G26 is twisted out of the helix and points down in to the major groove of helix 1, while the other three nucleotides form stacking interactions with the helical nucleotides. A24 is positioned intercalatively within the helical structure and introduces a slight kink in D5 (Figure 25b). Furthermore, the X-ray study describes the catalytic triad to be dynamic and somewhat disordered, but this is not supported by the NMR data. Additional structural information on similar folds come from a solution NMR-structure of D5 from *Pylaiella littoralis*[198] and an NMR structure of an extended U6 ISL found in the spliceosome.[199] All these structures share the same global architecture but with important local structural differences, especially in the internal bulge, and together confirm three magnesium ion binding sites at the tetraloop, the bulge, and the triad region.

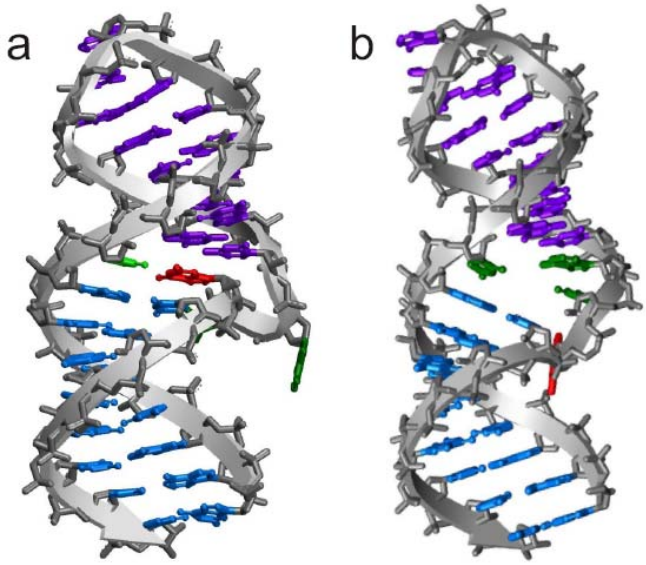


Figure 25. Structure of D5 from *Sc*.ai5γ as described by **a** crystallographic studies [122] and **b** NMR experiments [49]. Helix I (nucleotides G1 to A8 and U27 to C34) is shown in blue, nucleotides of helix 2 (G10 to C23), bulge nucleotides U9, A24 and C25 in green and G26 in red.

The large deviations in stacking properties of A24 in the dinucleotide bulge - which represents one of the proposed metal binding sites - led to the idea to study the structural changes in D5 from *Sc*.ai5γ dependent on the metal ion concentration. A24 was replaced by 2-aminopurine (2AP) for detection of local

conformational changes upon metal ion binding by using the fluorescence properties of 2AP as an additional technique to gather complementary structural and dynamic information. A range of metal ions were tested for their binding abilities to the bulge region in D5, i.e. Mg^{2+}, Ca^{2+}, Mn^{2+}, and Cd^{2+}. Mg^{2+} is freely available in any cell and considered to be the natural cofactor for group II introns.[44] Therefore, Mg^{2+} was used as the reference ion. Ca^{2+} neighbors Mg^{2+} sitting below in the periodic table of the elements and is also present in millimolar concentrations within certain compartments of cells.[200] This ion is found to inhibit the cleavage reaction of group II introns already in small concentrations.[201] The nature of this inhibition is unknown and no information is yet available on whether this influence is due to impaired folding or direct interference with catalysis. Aside Ca^{2+}, we have also looked at Mn^{2+} and Cd^{2+}, which are further M^{2+} ions important for RNA biochemistry.[41] Mn^{2+} has been described to accelerate the catalytic activity of some ribozymes [202-204] and is used as a probe in EPR [205] and NMR [206] experiments. Cd^{2+} is a thiophilic metal ion and is frequently used as a mimic of Mg^{2+} in thio-rescue studies.[193, 207]

Furthermore, the influence on the structure of D5 by the substitution of A24 with 2AP and the effect of metal ion binding to D5 was characterized in UV-melting experiments. Repeated transitions from folded and unfolded RNA structures and back assess the thermal stability of the molecules and report on changes in tertiary structure formation. Next to this, the melting experiments also measure the thermodynamic changes caused by the replacement of adenine with 2AP in the bulge or binding of metal ion cofactors.

3.2 Results

3.2.1 Two-piece D5: Base pairing of D5-17 with D5-19-2AP

Initially, two separate oligonucleotide strands with lengths of 17 and 19 nucleotides were used to reproduce the helical pattern and the bulge region of D5. Together, the strands contain the exact sequence of D5, which is self-complementary in most parts and allows the strands to base pair in the correct fashion. The original stem-loop is cut symmetrically in the middle of the tetraloop to give rise to the two halves of different length, D5-17 and D5-19 (Figure 26). This two-piece D5 was found to be highly reactive and to catalyze 5'-splice-site hydrolysis when added to an RNA construct comprising D123 and exon nucleotides (exD123), indicating no loss in function of the full ribozyme.[208] For an increased stability of the double-strand helix and increased transcription yields, the strands were extended by one nucleotide each to result in an extra G-C Watson-Crick basepair. The fluorescent nucleotide 2AP was incorporated in D5-19 at position 8 and the strand was renamed D5-19-2AP.

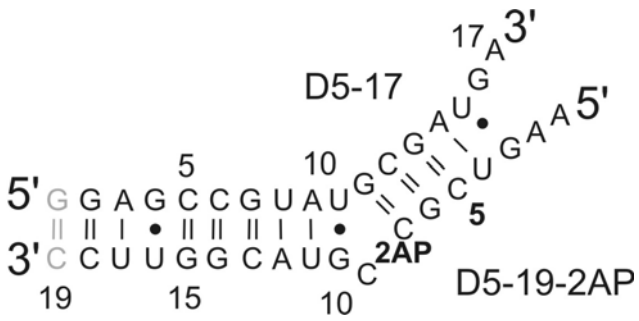


Figure 26. Depiction of the secondary structure of a two-piece D5 formed by base pairing of D5-17 and D5-19-2AP. The position of 2AP in D5-19-2AP is indicated in bold and the additional G-C basepair is indicated in gray.

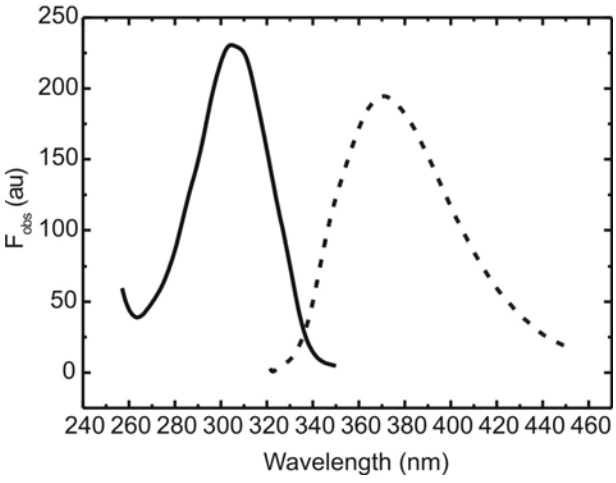


Figure 27. Excitation (solid line) and emission (dashed line) spectra of 2AP incorporated in D5-19-2AP (10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C).

Excitation and emission spectra of 2AP incorporated in D5-19-2AP (10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C) show a single band each with maximal intensities at wavelengths of 305 nm and 370 nm for excitation and emission, respectively (Figure 27).

Successive addition of concentrated amounts of either Mg^{2+}, Ca^{2+}, Mn^{2+} or Cd^{2+} to the two-piece D5 consistently leads to an initial increase in 2AP fluorescence at 370 nm, followed by a distinct decrease with linear tendency (Figure 28a). Mg^{2+}, Ca^{2+} and Cd^{2+} thereby show a larger absolute change in fluorescence than what is found for Mn^{2+}. Additionally, the transition from increase to linear decrease starts at lower concentrations of

Mn^{2+} than is the case for the other three metal ions. The curves could not be fitted with an equation describing the binding of one metal ion to one RNA molecule. Instead, this single site binding fit needed to be combined with a linear decrease to fully simulate the progressions (eq. 7, Section 5.4.2). Dissociation constants from these initial raw experiments turn out to be $K_{D,Mg}$ = 0.96 ± 0.20 mM, $K_{D,Ca}$ =0.56 ±0.09 mM, $K_{D,Mn}$ =0.07 ± 0.01 mM and $K_{D,Cd}$ =0.33 ± 0.05 mM (Table 2).

Table 2. Comparison of K_D values (μM) for Mg^{2+}, Ca^{2+}, Mn^{2+} and Cd^{2+} binding to the two-piece D5 and D5-36-2AP in various experiments. K_D values are the weighted mean from at least three experiments and error limits correspond to one standard deviation.

	DTT	HPLC	$K_{D,Mg}$	$K_{D,Ca}$	$K_{D,Mn}$	$K_{D,Cd}$
Two-piece D5	-	-	959 ± 198	566 ± 88	64 ± 10	325 ± 52
D5-36-2AP	-	-	12.3 ± 0.2	10.3 ± 0.2	4.3 ± 0.2	3.1 ± 0.2
	+	-	7.9 ± 0.2	9.3 ± 0.3	2.1 ± 0.1	n.d.[a]
	+	+	196.0 ± 9.9	144.1 ± 10.2	35.1 ± 3.2	n.d.[a]

[a] could not be determined because of binding of Cd^{2+} to DTT and impairment of 2AP fluorescence.

Metal ion titrations of D5-19-2AP only were carried out analogously and result in a fluorescent decrease after some initial minimal rise (Inlet Figure 28a). Single strand D5-19-2AP is assumed to be a random coil RNA without any specific structural organization in the presence or absence of metal ions. Therefore, titration experiments of D5-19-2AP indicate that the decrease does not reflect changes related to the bulge structure but is due to a structure-unspecific, systematic event. In an attempt to compensate for the linear decrease in two-piece D5, single strand D5-19-2AP titrations were used as baseline corrections. Subtraction of single strand D5-19-2AP fluorescence, normalized via division by the fluorescence value in the absence of metal ions, from the original duplex titration normalized in the same way yields curves, which no longer can be simulated by the single site binding isotherm to satisfying extent (Figure 28b) or meaningful fit values (data not shown).

Judging from the signal progressions during the titrations (Figure 28) without relying on any values from fits, addition of Mg^{2+} and Ca^{2+} to the two-piece D5 causes a similar change in 2AP signal and a comparable overall steepness of the increase, before and also after correction. The impact of Cd^{2+}, which induces a similar fluorescence change as Mg^{2+} and Ca^{2+} in duplex titrations, is reduced to smaller amounts and becomes comparable to what is found for Mn^{2+} after correction. In all titrations the decrease at higher concentrations after correction is to the most part, but not completely removed. Especially in the case for Mn^{2+} a considerable amount of decrease persists.

All in all, the correction attempt did not lead to any useful interpretation of the data and this initial approach was found to be not fully satisfying.

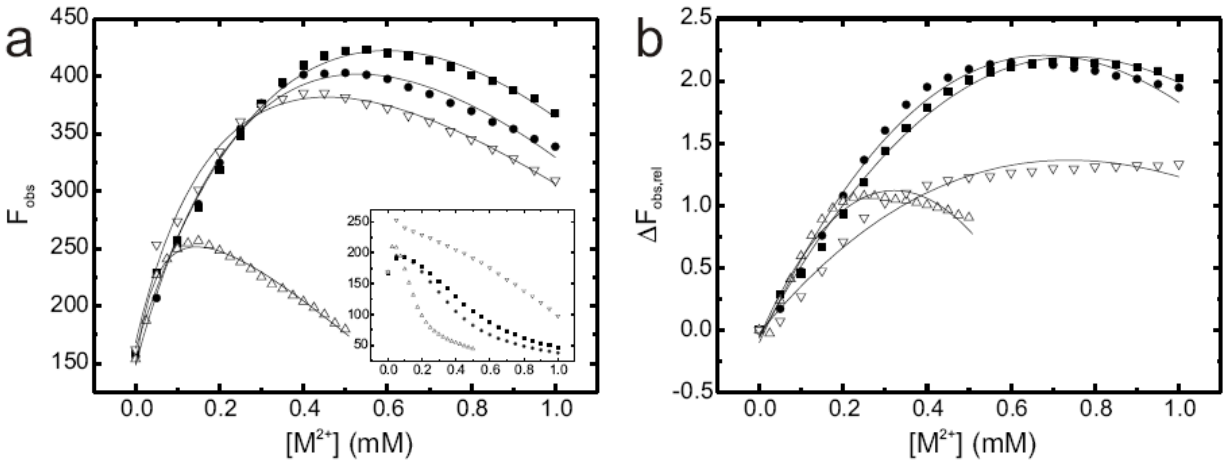


Figure 28. 2AP fluorescence intensity changes in D5-19-2AP upon metal ion addition. **a** Representative curves of two-piece D5 titrations. Inlet: D5-19-2AP single strand titrations. **b** Titration corrected as described in text. ■, Mg^{2+}; ●, Ca^{2+}; △, Mn^{2+}; ▽, Cd^{2+}. RNA concentration was kept at 10 µM. The solid lines represent the fit of the experimental data with the single site binding isotherm combined with a linear decrease (eq. 7, Section 5.4.2). λ_{ex} = 310 nm, λ_{em} = 370 nm, 10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C

3.2.2 Full-length D5-36-2AP

A full length construct D5-36-2AP (extended with an additional G-C base pair at the helix end) containing 2AP at position A25 (Figure 29) was studied in subsequent experiments in an attempt to simplify the results and avoid possible saturation problems during experiments involving two separate strands. Additionally, an advanced titration technique developed in the Walter Lab, University of Michigan, Ann Arbor, USA was applied to balance RNA dilution effects during the metal ion additions (see Section 5.4.2).[209] The system was further allowed to reach equilibrium conditions for a time interval of 4 min after every metal ion addition and concomitant mixing before data collection. Fluorescence spectra of D5-36-2AP were comparable to Figure 27 with maximal bands at 305 and 370 nm for excitation and emission, respectively.

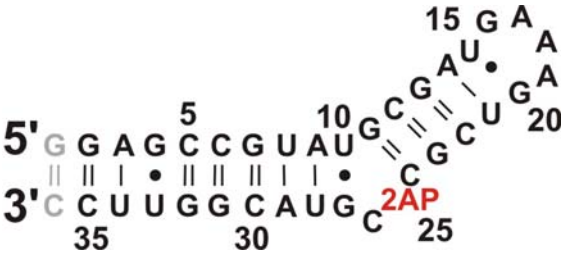


Figure 29. Secondary structure of D5-36-2AP. Position of the fluorescent nucleotide 2AP is highlighted in red.

Comparable to the previous experiments with two separate strands, full-length D5-36-2AP shows an initial increase in 2AP fluorescence for all metal ion tested (Figure 30a). Although reduced, a decrease in fluorescence at higher metal concentrations is still measured.

The decrease varies in every individual experiment and independently of the metal ion identity. To correctly account for the decrease in the experiments and to end up with accurate estimations of K_D, several experiments spanning a range of metal ion concentrations up to 800 µM were performed. Representative examples for Mg^{2+}, Ca^{2+}, Mn^{2+}, and Cd^{2+} are shown in Figure 30a. Addition of Mg^{2+} and Ca^{2+} leads to a twofold maximal increase in fluorescence intensities, whereas Cd^{2+} induces smaller changes in fluorescence levels. Mn^{2+} seems to induce fluorescence amplitudes comparable to Mg^{2+} and Ca^{2+}, but the titration is complicated by the overlay of a large decrease. Strikingly, the dissociation constants are roughly hundredfold reduced compared to the values found in titrations with the two-piece D5 set-up: $K_{D,Mg}$ = 12.3 ± 0.2 µM, $K_{D,Ca}$ = 10.3 ± 0.2 µM, $K_{D,Mn}$ = 4.3 ± 0.2 µM and $K_{D,Cd}$ = 3.1 ± 0.2 µM (Table 2).

Although at a reduced level, the linear decrease remains present in all experiments after the initial increase. To further avoid this decrease, we tested for the effect of oxygen scavengers. Indeed, addition of DTT results in the desired effect, which is most obvious in case of Mn^{2+} (Figure 30b)!! Slight remnants of the decrease can still be seen in a very few experiments, and sometimes even a slight increasing tendency is detected. Again, Mg^{2+} and Ca^{2+} are found to behave in similar ways, whereas Mn^{2+} does not induce the same maximal fluorescence change also in the presence of DTT. Dissociation constants were found to be $K_{D,Mg}$ = 7.9 ± 0.2 µM, $K_{D,Ca}$ = 9.3 ± 0.3 µM and $K_{D,Mn}$ = 2.1 ± 0.1 µM (Table 2). Titrations of D5-2AP with Cd^{2+} in buffer containing DTT were not possible, since Cd^{2+} is thiophilic and binds to the thio-groups on DTT rather than the RNA and this competition considerably impairs 2AP fluorescence.

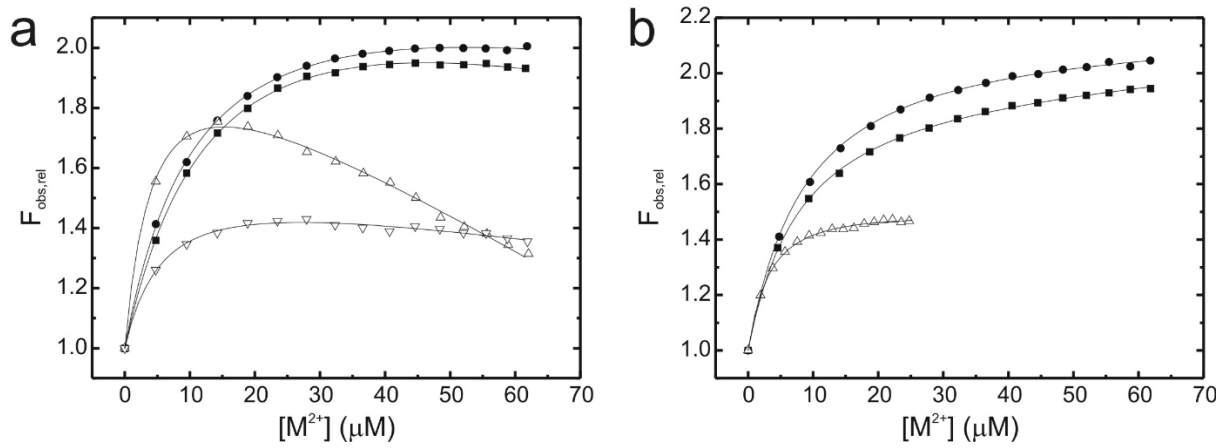


Figure 30. Relative change in 2AP fluorescence in D5-36-2AP upon metal ion addition **a** in buffer containing no DTT and **b** in buffer containing 25 mM DTT as a ROS scavenger. ■, Mg^{2+}; ●, Ca^{2+}; △, Mn^{2+}; ▽, Cd^{2+}. The solid lines represent the fit of the experimental data with the single site binding isotherm combined with a linear term (eq. 7, section 5.4.2). λ_{ex} = 310 nm, λ_{em} = 370 nm, 10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C

Experiments under optimized conditions

For a last series of experiments D5-36-2AP was purified by HPLC and titrated with Mg^{2+} and Ca^{2+} in a freshly prepared buffer (10 mM MOPS, pH 6.0, 10 mM KCl) containing 25 mM DTT. The relative fluorescence maximally increases 3.5 fold and the K_D values change to higher numbers compared to D5-36-2AP titrations before HPLC-purification (Figure 31, Table 2). Otherwise, the behavior of 2AP fluorescence upon metal addition remains comparable to earlier experiments and is fitted with a single site binding equation assuming no linear decrease (eq. 7, a = 0). The fits converge better if the initial measurement point at 0 mM is omitted and result in dissociation constants of $K_{D,Mg}$ = 196.0 ± 9.9 µM, $K_{D,Ca}$ = 144.1 ± 10.2 µM and $K_{D,Mn}$ = 35.1 ± 3.2 µM (Table 2). Mg^{2+} and Ca^{2+} induce similar fluorescence changes with comparable K_D values, whereas binding of Mn^{2+} is described by a higher affinity but results in reduced maximal 2AP fluorescence change.

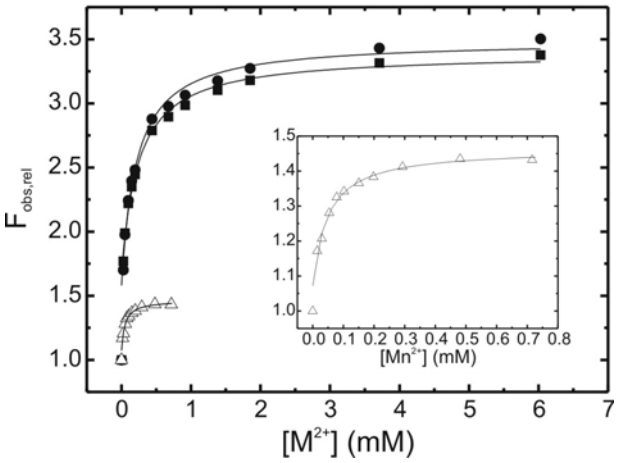


Figure 31. Relative change in 2AP fluorescence of HPLC-purified D5-36-2AP upon addition of Mg^{2+}(■), Ca^{2+} (●) and Mn^{2+} (△). Intlet: enlargement of titration with Mn^{2+}. Fits of the experimental data (solid lides) describe a single site binding isotherm (eq. 7, a = 0). λ_{ex} = 310 nm, λ_{em} = 370 nm, 10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C

3.2.3 Control experiments

Although 2AP is being widely employed as a reporter of structure and dynamics of nucleic acids,[64, 210, 211] the exact fluorescence quenching mechanism and its dependence on the sequence and the structural context is complex and only understood in outlines.[137, 212] Usually, the results are interpreted in qualitative amounts and need to be verified and validated in control experiments to rule out influences other than RNA structural changes that possibly lead to fluorescence quenching. Additionally, the persistent appearance of a linear decrease in experiments suggests a systematic loss in fluorescence, the origin of which was initially unknown and tested for in a number of control experiments.

2AP is photochemically stable during experiments

D5-36-2AP, D5-19-2AP as well as isolated 2AP in context of the cyclic monophosphate (2-NH_2-PuMP, Figure 32a) were continuously (Figure 33a) or repetitively excited (Figure 11a, points 0-10). In all experiments, only a minimal long-term loss in fluorescence up to 1 hour is recorded, confirming satisfactory photostability of 2AP during our experiments.

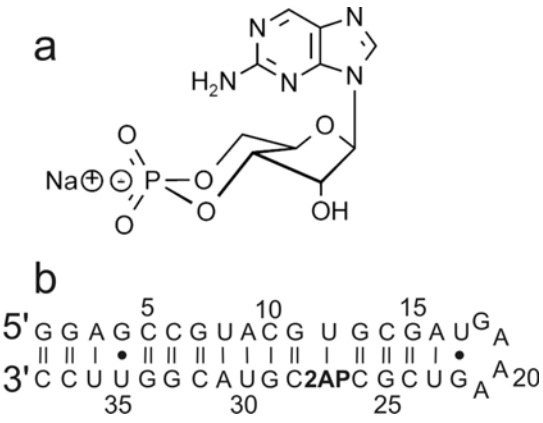


Figure 32. Molecules used to report on 2AP base stacking. **a** Secondary structure depiction of D5-38-2AP. 2AP is indicated in bold. **b** Chemical drawing of 2 aminopurine riboside-3'-5'-monophosphate (2-NH_2-PuMP), sodium salt.

No structural change in D5-36-2AP is observed upon addition of monovalent cations

Stepwise addition of the monovalent cation K^+ in place of divalent cations does not result in an increase of 2AP fluorescence in D5-36-2AP up to a concentration of 180 mM KCl (Figure 33b). Therefore, the monovalent cation K^+ cannot induce a local structural change in the bulge region even at elevated concentrations. These results clearly indicate the specific involvement of a metal ion of higher valency at a binding site in the bulge of D5. Instead, the ubiquitous decrease is recorded. This decrease therefore does not depend on the valency of the metal ion used.

Base stacking interactions reduce 2AP fluorescence

Fluorescence of 2AP is known to be quenched by stacking interactions.[137] 2AP in a perfect double helix is fully stacked and fluorescence is quenched. If 2AP is placed in loops or bulges, where stacking interactions are smaller than in the double helix, higher fluorescence

of 2AP is measured. To test this, 2AP fluorescence in context of an A-form helix and influence of metal ion addition was measured on a construct D5-38-2AP, in which the originally bulged nucleotides A25 and C26 in D5-36-2AP are matched by their Watson-Crick partners to result in a perfect hairpin (Figure 32b). Indeed, measured absolute intensities are reduced 3-4 fold compared to intensities in D5-36-2AP and no change in fluorescence apart from the linear decrease is detected upon Mg^{2+} addition (Figure 33c). Therefore, 2AP in the bulge of D5-36-2AP reports on a partially unstacked nucleotide and the fluorescence decrease is not associated to degree of stacking interactions.

Metal ions do not influence 2AP fluorescence directly

A 2-aminopurine riboside-3'-5'-mono-phosphate (2-NH_2-PuMP, a cyclic fluorescent nucleotide, Figure 32a) was used to simulate the behavior of 2AP in the absence of neighboring bases and to further check for direct quenching of 2AP fluorescence by metal ions. Titration experiments were performed with Ca^{2+} and Cd^{2+} as well as buffer for comparison. Absolute fluorescence intensities of 2-NH_2-PuMP are very high compared to 2AP in context of RNA polymer like D5-36-2AP, as is expected for a maximally unstacked nucleotide with full quantum yield of fluorescence. A constant decrease in 2AP fluorescence occurs in all titrations, which was independent of the nature of metal ion (Figure 33d), metal ion concentration or the presence of metal ions altogether (data not shown). Addition of buffer only resulted in the same relative decrease, proving no direct 2AP quenching by metal ions.

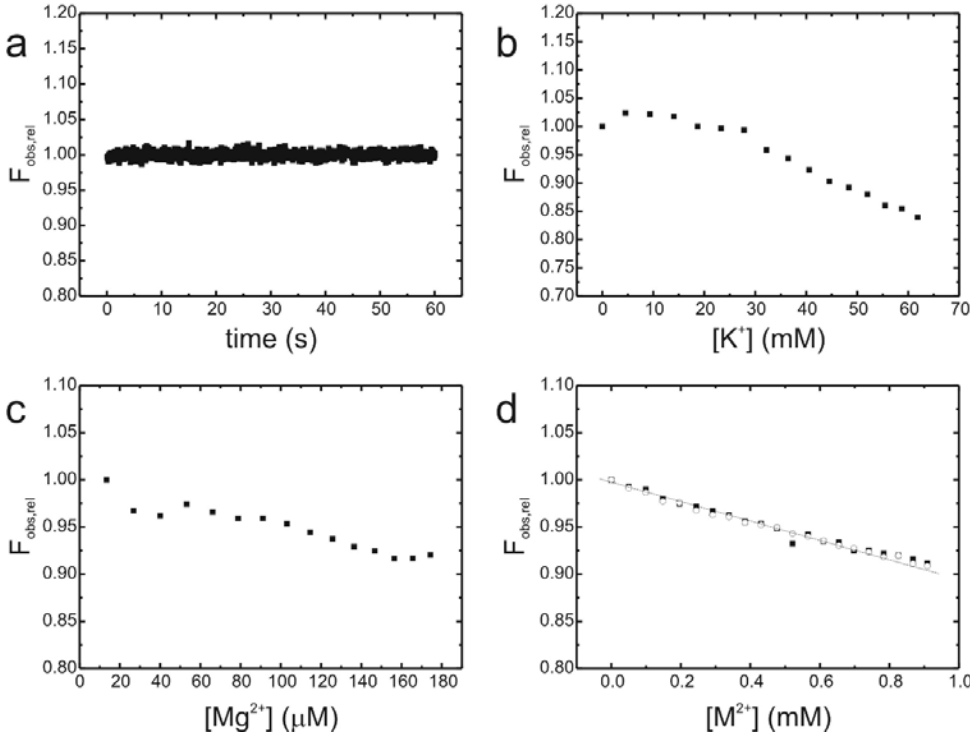


Figure 33. 2AP control experiments. **a** 2-NH_2-PuMP fluorescence intensity during constant excitation for 1h. **b** Fluorescence of D5-36-2AP upon addition of K^+. **c** 2AP signal from D5-38-2AP in titration with Mg^{2+}. **d** Influence of metal ion addition on 2-NH_2-PuMP fluorescence. ■, Ca^{2+} and ○, Cd^{2+}. Solid line depicts decrease in fluorescence measured upon dilution with buffer only. λ_{ex} = 310 nm, λ_{em} = 370 nm, 10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C

Oxygenation of solution quenches 2AP fluorescence

The decrease in fluorescence, which was omnipresent in all experiment so far and the cause of which could not be assigned for a long time, was eventually found to depend on the manual pipeting procedure. If the solution is left undisturbed in between measurements, only a minimal change of fluorescence is detected. Repeated mixing of the solution in between measurements instead causes a considerable loss of fluorescence (Figure 34a). 30 rounds of pipeting the sample up and down to mix contents leads to a substantial linear decrease. The slope of the decrease is reduced if the solution is only mixed 10 times in between measurements. Increased contact of the solution with air in combination with excitation by light causes increased levels of reactive oxygen species (ROS) that react with the excited states of 2AP and quench its fluorescence. Therefore, mixing was reduced to a minimum during titration experiments to minimize the fluorescence decrease that is of non-structure related manner.

Oxygen scavengers protect D5-36-2AP from loss of fluorescence

To further circumvent this problem, the samples were degassed before the experiments and oxygen scavenger molecules, i.e. DTT; TCEP or Vitamin C were added to the solution. TCEP is found to interact with electronic structure of 2AP, since the 2AP fluorescence in the presence of TCEP is dramatically reduced and excitation as well as emission spectra change considerably. Addition of Vitamin C resulted in similar findings. Therefore, both TCEP and Vitamin C were found to be not suitable for fluorescence experiments with RNA.

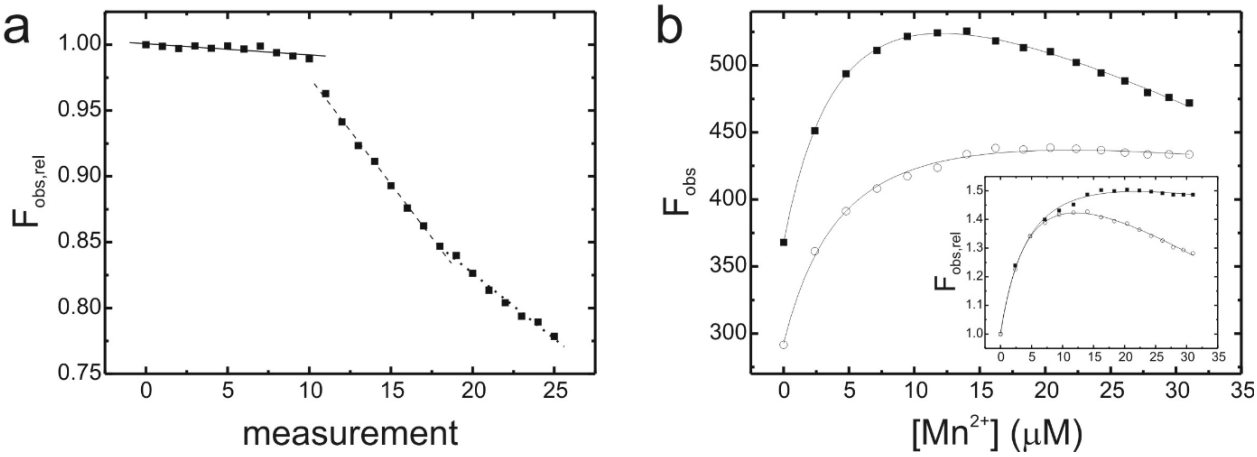


Figure 34. Influence of oxygenation and DTT addition on 2AP fluorescence. **a** Decrease in fluorescence is dependent on the mixing procedure. Fluorescence stays nearly constant in repeated measurements without disturbing the solution (measurements 0-10, solid linear fit), 30 rounds of pipeting the solution up and down in between the scans introduce considerable decay (measurements 11-18, dashed linear fit), whereas 10 rounds of mixing before the measurement lessen the amount of decrease (measurements 19-25, dotted linear fit). **b** D5-36-2AP titration with Mn^{2+} in the presence (■) and the absence (○) of DTT. Inlet: comparison, fluorescence signal is normalized via division by the value recorded in the absence of divalent metal ions. λ_{ex} = 310 nm, λ_{em} = 370 nm, 10 mM MOPS, pH 6.0, 10 mM KCl, 25 °C

DTT on the other hand, which together with β-mercaptoethanol is generally used in fluorescence measurements,[213, 214] does not influence the positions or shapes of excitation and emission bands in the spectra. DTT proved to be the most efficient oxygen scavenger and led to the most promising results. In the presence of DTT, the linear decrease observed in all experiments is reduced to a minimum. This is most obvious in the case of Mn^{2+}, where the loss in fluorescence was previously found to be most pronounced (Figure 34b). The absolute amount of fluorescence is slightly reduced in the presence of DTT. Metal ion addition still induces an increase, which turns out to be almost identical to the former experiments in the absence of DTT if fluorescence signals are normalized via division by the value in the absence of divalent metal ions (Inlet Figure 34b). One drawback of the utilization of DTT is the fact that DTT contains thio groups. Since Cd^{2+} is thiophilic and forms complexes with thio groups in DTT, it was not possible to perform Cd^{2+} titration experiments in the presence of DTT. Excitation and emission spectra were impaired and titration experiments led to no meaningful results.

3.2.4 Thermal stability changes in D5 upon introduction of 2AP and titration with metal ions

The specific binding of divalent metal ions to D5 has been proposed before.[49, 122, 175] The binding of cofactors potentially induces structural rearrangements and is reflected in changes in thermal stability. Furthermore, possible structural aberrations due to the identity of the metal ion cofactor (differences in binding positions, coordination, orientation or size of metal ion cofactor) can also be detected by measuring UV-melting curves. Melting of the RNA in these experiments is recorded in an aqueous solution and in the presence of 100 mM KCl. Under these unbuffered (i.e. in the absence of MOPS) and exclusively monovalent metal

Table 3. Melting temperatures T_m (°C) and metal ion dependency for thermal melting for D5-36-2AP and D5. K_D (mM) are weighted mean values from three heating cycles and the error limits correspond to one standard deviation.

	MOPS	[KCl]	[M^{2+}]	$T_{m,K}$	$T_{m,Mg}$	$T_{m,Ca}$	$K_{D,Mg}$	$K_{D,Ca}$
D5-36-2AP	+	10	0	55.5 ± 0.4	-	-	-	-
	+	10	1	-	80.8 ± 0.5	-	-	-
D5	+	10	0	58.9 ± 0.3	-	-	-	-
	+	10	1	-	83.3 ± 0.3	-	-	-
	-	100	0	73.1 ± 1.2	-	-	-	-
	-	100	20	-	88.1 ± 0.9	83.2 ± 0.6	-	-
	-	100	-	-	-	-	1.1 ± 0.2	1.2 ± 0.2

ion conditions, the RNA melts in a single transition described by T_m = 73.1 ± 1.2 °C (Figure 35a, b). Thermal stability changes in D5 are measured upon metal ion addition from 0-20 mM Mg^{2+} or Ca^{2+}. In the presence of divalent metal ions, the formerly linear baseline drift converts into an additional exponential transition. This first transition at lower temperatures does not experience any stabilizing effect upon metal ion addition, it is found to scatter around 65-75 °C in an undefined manner. On the other hand, the presence of both Mg^{2+} and Ca^{2+} results in an increase in melting temperature of the second transition at higher temperature (Figure 35a). The stabilizing effect is found to be more efficient for Mg^{2+}. Maximal transition temperature are 88.1 ± 0.9 °C at 20 mM Mg^{2+} and 83.2 ± 0.6 °C at 20 mM Ca^{2+} and showed a difference of about 5°C. The apparent K_D values are $K_{D,Mg}$ = 1.1 ± 0.2 mM and $K_{D,Ca}$ = 1.2 ± 0.2 mM if the experimental data are fitted with the Hill equation (eq.11).

In order to estimate the influence of 2AP substitution on the stem-loop structure stability, UV melting curves of D5-36-2AP and the analogous RNA construct D5 without 2AP were measured in a buffer containing 10 mM MOPS, pH 6.0 and 10 mM KCl. Upon gradual heating from 20-95°C, the RNA melts in a single transition at T_m of 58.3 ± 0.3 °C or 55.5 ± 0.4 °C for D5 and D5-36-2AP, respectively (Figure 35b) if assuming a linear baseline drift in the folded state (eq. 8). In the presence of 1 mM Mg^{2+}, a large change in the melting profile can be seen (Figure 35b). The melting profile is shifted to higher temperatures, indicating extensive stabilization by Mg^{2+}, and the formerly linear tendency at lower temperatures is reduced to a minimum. Instead, a double melting transition appears. The lower transition is not stable in repeated heating cycles and clusters around 74 °C and 69 °C for D5 and D5-36-2AP, respectively. In contrast, the second transition at higher temperature is uniformly found

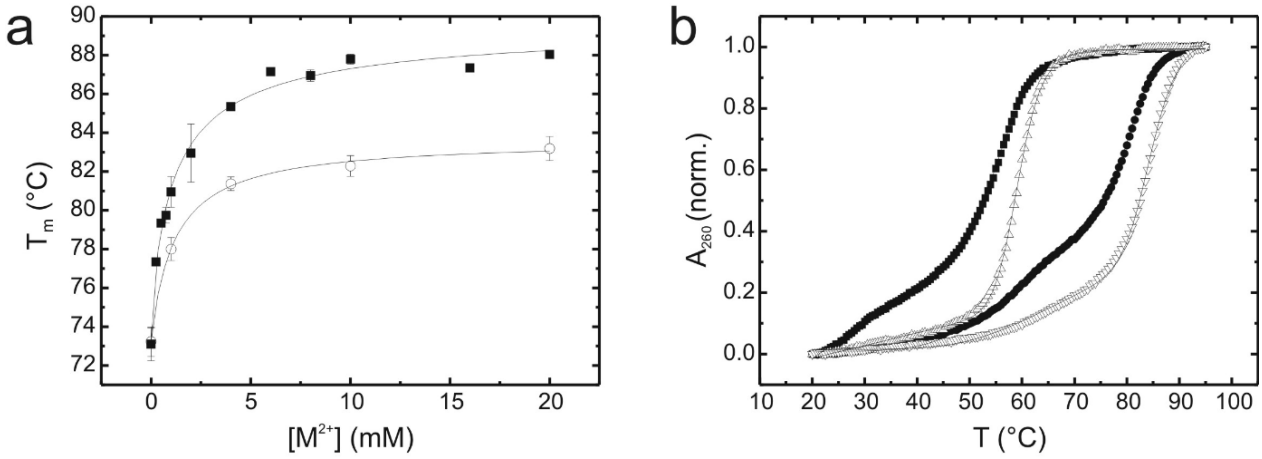


Figure 35. Influence of metal ions and introduction of 2AP on the melting transition of D5. **a** Increase in melting temperature of D5 upon metal ion addition. ■, Mg^{2+}; ○, Ca^{2+}. Solid lines depict fit with Hill equation (eq. 11) and reveal apparent K_D values of $K_{D,Mg}$ = 1.1 ± 0.2 mM and $K_{D,Ca}$ = 1.2 ± 0.2 mM. **b** Melting curves of D5 (open symbols) and D5-36-2AP (full symbols) in the absence and the presence of 1 mM Mg^{2+}. Solid lines are fits describing one (0 mM Mg^{2+}, eq. 8) or two (1 mM Mg^{2+}, eq. 9) melting transitions. ■, D5-36-2AP, 0 mM Mg^{2+}; ●, D5-36-2AP, 1 mM Mg^{2+}; △, D5, 0 mM Mg^{2+}; ▽, D5, 1 mM Mg^{2+}

at 83.3 ± 0.3 °C and 80.8 ± 0.5 °C for D5 and D5-36-2AP, respectively. This transition temperature of the melting event at high temperatures consistently shows a reduction of 3 °C upon 2AP introduction into D5, both in the presence and the absence of Mg^{2+}.

3.3 Discussion and Conclusion

3.3.1 Unstacking of 2AP in the D5 bulge upon metal ion binding

Addition of divalent metal ions to D5-36-2AP results in an increase in fluorescence of 2AP positioned at the dinucleotide bulge (Figures 30 and 31). This signal change is thought to be due to alterations in the local conformation caused by the binding of a metal ion to or in close vicinity of 2AP. Compared to D5-36-2AP, the fluorescence intensity of the cyclic monophosphate derivative 2'-NH_2-cPuMP is found to be dramatically elevated, whereas the perfect hairpin D5-38-2AP shows a >3-fold reduced signal under identical conditions. This observation beautifully reflects the influence of base stacking on 2AP fluorescence, showing maximal fluorescence in the absence of neighboring bases and indicating extensive quenching in context of an A-form helix.

Intermediate fluorescence of 2AP at position 25 in the dinucleotide bulge of D5-36-2AP indicates a partially unstacked nucleotide. Upon divalent metal ion binding in vicinity of this nucleotide, changes in interactions with surrounding bases occur and result in loosening of the bulge structure as reported by increasing fluorescence signals. A25 apparently experiences reduced stacking interactions in the partially helical environment and becomes more exposed to solution. The extent, to which this is achieved, cannot be estimated in numbers, since the experiments conducted in this work only result in a qualitative description of the system. However, the results indicate that metal ion binding to the bulge induces a possible rotation of a bulge nucleotide around the RNA backbone out of the helix content.

It is not possible to verify either the X-ray[122] or the NMR[49] structure (Figure 25), since both depict a partially unstacked A25 and furthermore deviate in buffer and cation conditions. For the NMR structure D5, was lyophilized from a buffer consisting of 100 mM KCl, 10 µM EDTA at pH 6.5 and then dissolved in H_2O/D_2O for experiments. Crystallization was successful using the sitting-drop vapor diffusion technique with an RNA-construct comprising domains 5 and 6 in buffer conditions of 10 mM HEPES-KOH (pH 7), 2.5 mM $MgCl_2$ and reservoir solution of 50 mM HEPES-KOH (pH 7), 10 mM $MgCl_2$, 1-1.2 M $(NH_4)_2SO_4$ and 0.1 mM spermine.[122] The bulged nucleotides in the X-ray structure form intermolecular lattice contacts to sister molecules, which might influence the conformation adopted in the crystal. As shown in the 2AP titration experiments, the positions of the bulged nucleotides also change with metal ion conditions and cause deviations in the reported structures.

3.3.2 Decreasing fluorescence drift during measurements

All experiments are complicated by the appearance of a linear decrease in fluorescence signal occurring throughout the titrations. The nature of this loss of fluorescence was hidden to us for a long time and suspected to be of systematic and non-structure-related origin. The linear decrease could not be described by a value that was constant throughout different experiments as it varies in every individual experiment. Control experiments to test for direct quenching of 2AP fluorescence by metal ions were carried out with a cylic 2AP monophosphate (2'-NH_2-cPuMP). Addition of Ca^{2+} or Cd^{2+} over a large concentration range revealed no additional quenching compared to dilution with buffer only. Therefore, addition of metal ions *per se* does not result in fluorescence quenching of 2AP. Dilution effects certainly lower the measured intensities. However, they can be measured as well as calculated and turn out not to fully account for the total loss of fluorescence. A gradual loss of fluorescence during constant or repeated excitation due to photobleaching is detected, but is determined to be minimal.

The loss of fluorescence turned out to distinctively depend on the mixing procedure in between measurements (Figure 34a). Reduced and careful manual mixing, during which the amount of bubble generation in solution was kept to a minimum, results in a much less pronounced reduction in fluorescence. Molecular oxygen and its ROS are known to be prominent and efficient quenchers of both singlet and triplet excited states. Increased contact with oxygen from air causes oxygenation of the solution. Subsequent excitation with light leads to an elevated level of ROS that can react with excited states to ultimately result in fluorescence quenching. Therefore, mixing in between data collection was restricted to a minimum of 3 rounds of pipeting up and down. Additionally, a ROS scavenger in form of DTT was added to the solution, which results in near elimination of the linear decay (Figure 34b). Complete prevention of the linear drift could not be achieved, which might be associated in part to concentration, age and decomposition state of DTT.

3.3.3 Influence of metal ion identity

Independent of the experimental set-up, the behavior and the amplitude change of 2AP fluorescence upon addition of Mg^{2+} and Ca^{2+} remained similar (Figures 28, 30 and 31). Therefore, the impact of binding of Mg^{2+} or Ca^{2+} on the local structure in the dinucleotide bulge is comparable, both in absolute and relative changes in fluorescence as well as in the values of the dissociation constant K_D (Table 2)

Mn^{2+} and Cd^{2+} on the other hand were found to deviate from the two metal ions mentioned before and show smaller K_D values as well as reduced changes in intensity. Buffer conditions and correction procedures influence titrations with Mn^{2+} and Cd^{2+} much more and the general results set Mn^{2+} and Cd^{2+} apart from Mg^{2+} and Ca^{2+}. The higher affinities of Mn^{2+} and Cd^{2+} but their smaller influence on local conformation indicate altered binding to and different local structure of D5 compared to Mg^{2+} and Ca^{2+}.

The influence of oxygenation detected by the decrease in fluorescence was most pronounced in the case of Mn^{2+} (Figures 28 and 30). It seems that the interaction of this particular metal ion with 2AP renders the excited states most susceptible to react with oxygen species. The manifold possible mechanisms that influence 2AP fluorescence make it difficult to deduce the reason for this. It could be an alteration in charge-transfer processes of 2AP due to the binding of metal ions or modifications in π-stacking interactions due to a slightly altered conformation of the bulge in the presence of Mn^{2+} compared to the other metal ions, just to mention two possibilities.[212] However, it is clear that the electronic structure of 2AP in the excited state in context of the RNA is not the same with every metal ion tested. The scavenger effect of DTT was most efficient in the case of Mn^{2+} (Figure 34b). Titrations with Cd^{2+} in the presence of DTT were not possible since Cd^{2+} binds with high affinity to thiogroups in DTT rather than acting as a scavenger for ROS, leaving RNA free for oxidative damage and, in addition, severely derange 2AP excitation and emission spectra.

3.3.4 Dissociation constants of metal ion binding to D5

Formal ligation of the initially separated halves of D5 led to a nearly hundredfold reduction in K_D for metal ion binding to the D5 bulge as detected by 2AP fluorescence (Table 2). The discontinuity of the tetraloop results in inter- rather than intramolecular formation of hydrogen bonding and apparently reduces the metal ion binding affinities of full-length D5. The tetraloop itself is described to be a metal binding site[49, 175] and might guide the overall formation of the molecule, since the tetraloop introduces a turn and brings the two halves of D5 in close proximity and facilitates basepair formation. In this study, the focus is on the local structure of the dinucleotide bulge, which also provides a metal binding site.[49, 175] Intensity changes in 2AP in the bulge should not directly report on binding of a metal ion at the tetraloop, but if the absence of the tetraloop changes the overall metal ion requirements this certainly would also be reflected in the K_D for binding of metal ions to the dinucleotide bulge. Furthermore, the initial system is complicated by the use of two separate strands to constitute the full-length D5. Variations in concentrations of the separate strands influence the

binding ratio and full saturation of the 2AP labeled strand might not always be completely reached.

Addition of DTT as an oxygen scavenger results in quenched absolute fluorescence signals and leads to changes in K_D values (Figure 34, Table 2). Nevertheless, relative changes in fluorescence remained constant to a large extent and changes in K_D were relatively small (10-50 %). Differences in K_D might furthermore be connected to the fact that the linear decay overlays with the actual data and perfect resolution might still not be achieved. Overall and apart from Cd^{2+}, DTT addition to the buffer is a valuable alternative to avoid quenching of fluorescence that is unrelated to structural changes.

The K_D values vary strongly with different conditions (Table 2). The reproducibility is very good when working under constant conditions, i.e. identical buffer stocks and sample aliquots. HPLC analysis reveals a typical pattern of shorter fragments due to the chemical synthesis with the largest peak corresponding to the full-length product and a few minor peaks eluting at shorter retention times. The isolation of the full-length product also removes residual organic chemicals used during the chemical synthesis of D5-36-2AP that might influence 2AP fluorescence. An overall increase in fluorescence that could be fitted with the single site binding isotherm[25] was confirmed in all experiments. However, after HPLC-purification the fluorescence intensity changes were elevated and K_D values increased roughly 15 fold.

3.3.5 Melting curves

Melting temperatures of the D5 and D5-36-2AP hairpin structures are found to highly depend on monovalent, as well as divalent metal ion concentrations. The increased thermal stability indicates distinct metal ion binding and thereby formation and stabilization of higher order structures. A stabilization of 15 °C is found when increasing KCl from 10 mM (10 mM MOPS, pH 6.0) to 100 mM in H_2O (Table 3). Divalent metal ions stabilize D5 and D5-36-2AP even further. In a background of 10 mM KCl, the melting temperature is found to shift up to 25 °C by the addition of 1 mM $MgCl_2$. The maximal thermal stability of the D5 hairpin is found to depend on the identity of the divalent metal ion. At 20 mM Mg^{2+} the maximal melting transition is 5 °C higher than what in 20 mM Ca^{2+}, indicating an altered structural conformation of D5 depending on divalent ion identity (Table 3).

The UV-melting profiles in the absence of divalent metal ions show one single melting transition. The observed baseline drift could correspond to weak temperature-dependence of the folded hairpin state. But the slope of this drift is high compared to baseline drifts in

melting profiles for of other small RNAs[215, 216] and might possibly correspond to another melting event, which linearly depends on temperature in the absence of divalent metal ions. This same melting event experiences additional stabilization upon addition of divalent metal ions and becomes a sigmoidal transition, to which a melting temperature can be assigned. However, the linear baseline drift in the folded state is always included in the fitting equation (eqs. 8 and 9).[217] In the presence of divalent cations, the melting curves depict two melting events at two different temperatures. There are (at least) two possible mechanisms that explain this behavior.

First, the two transitions could indicate the presence of two subfragments of different stabilities[218] and would then be attributed to the two helices in D5 that are separated by the bulge. While the hairpin structure apparently melts cooperatively in the absence of divalent metal ions, the structures of the two helices become stabilized and denature distinguishably – although most probably not completely independently - upon heating in the presence of divalent metal ions. The lower melting temperature corresponding to the melting of the shorter helix 1 (5 basepairs, 3 G-C basepairs, Figure 29) is found to scatter between 60-77 °C, not revealing any distinct divalent metal ion dependencies. The second melting event at high temperatures on the other hand is a typical sigmoidal transition that shows distinct dependencies on the metal ion concentration. This transition describes the melting of the longer helix 2 (10 basepairs, 5 G-C basepairs), which is known to harbor a specific Mg^{2+} binding site at the AGC-triad.[49, 175, 193] It is most probably not only this metal ion binding site at the catalytic triad that is reflected by the apparent K_D of the temperature shift of the second transition, since the other metal binding sites at the bulge and the tetraloop will certainly influence the stability of the hairpin structure, too. Yet, UV absorbance increases upon destacking, described by the so called hypochromic effect, and the rise in temperature increases absorbance due to loss of stacking,[219] which is most pronounced in melting of helices compared to (already) single stranded structures like bulges and loops.

Alternatively, the two-transition melting profile could correspond to the loss of a metal ion that binds to the bulge and subsequent two-state break up of the hydrogen bonding network in the hairpin possibly starting from the bulge region, which is no longer stabilized by the metal ion. Compared to the thermal stability of the sum of all the hydrogen bonds in the hairpin ($\Delta\Delta G°_{HB}$ ~2-10 kJ/mol[220-222]), the gain in thermal energy of metal ion binding to RNA is weak, and therefore this interaction will be resolved first.

The latter mechanism seems to be more plausible to explain the observed melting profiles of D5. Hairpins are described to melt in a two-step transitions and have high melting

temperatures between 55-80 °C,[215, 223] even when the stem consists of as little as three to five basepairs or contain a mismatched basepair in the stem.[216, 224] The smallest RNA system described to unfold in at least two steps, is a hairpin with an internal loop of 6-10 nucleotides, but only if the two helices are of sufficient stabilities, i.e. containing a lot of G-C basepairs.[215, 218] The D5 molecule is a short hairpin structure with an small internal dinucleotide bulge and its tertiary structures depict a continuous A-form helix with only a slight kink.[49] It is hard to imagine that the two helices of D5 would melt independently or with melting temperatures that differ by 15-25 °C rather than overlay or melt simultaneously.

3.3.6 Conclusion

In this chapter, the focus is set on different aspects of the influence of different divalent metal ion that bind to D5. It was previously found that changing the divalent metal ion from Mg^{2+} to Ca^{2+} results in inhibition of the catalytic activity of the whole ribozyme.[201] It is unclear whether this finding is due to a replacement of a metal ion that is directly involved in catalysis or if it reflects an effect of changes in ribozyme (sub)structure(s) that lead to an inactive molecule. D5 is known to sit at the heart of the active site in group II intron ribozymes[109, 127] and therefore, inactivation of catalysis might be caused by alterations of the central D5 element. D5 is indispensable for catalysis[182, 183] and structural data on this small hairpin structure is available, too.[49, 122] It was previously found that the metal ion binding sites in D5 persist in the isolated as well as the *Sc.*D135 ribozyme, suggesting a similar structure of D5 in the absence and the presence of the ribozyme scaffold.[175] In an attempt to address the role of metal ions in function for folding or catalysis, isolated D5 was studied here with regard to structural changes upon the addition of different divalent metal ions.

The thermal stabilities of the D5 and the D5-36-2AP hairpins increase through binding of mono- and divalent metal ions. Maximal stability was reached by the use of divalent ions and was indeed found to depend on identity of divalent ion used: Mg^{2+}, which is the natural cofactor of the ribozyme, introduces a significantly larger stabilizing effect than Ca^{2+}. On the other hand, local structural changes at the bulge region – which is described to harbor a specific metal binding site and where deviations in structural data of X-ray and NMR studies are largest - were found to be comparable upon addition of Mg^{2+} and Ca^{2+}, both in the amount of fluorescence change and the dissociation constant K_D from fluorescence experiments.

Therefore, contribution of altered thermal stabilization of D5 might indicate a contribution of D5 to the inhibitory effect of small amounts of Ca^{2+} found on the catalytic

reaction of the whole ribozyme. But it is not the influence of specific binding of a metal ion to the nucleotide bulge that causes these differences, since fluorescence experiments that report on the binding of a metal ion to the bulge lead to similar results for Mg^{2+} and Ca^{2+}. Further experiments on metal ion binding to specific sites in D5 would involve the other two described binding sites, i.e. placing 2AP at positions in the tetraloop and in the AGC-triad. Towards this end, the thermodynamic stability studies might indicate a preference for an involvement of the AGC-triad.

K_D values obtained from fluorescence and thermal denaturation experiments do not coincide, but rather reveal differences of a factor of 5-8 (Table 4). The fluorescence experiments report on local binding of specific metal ion to the bulge due to placing 2AP in the bulge region of D5. Thermal stability measurements on the other hand might reflect metal ion binding to the AGC-triad or, more likely, the combination of divalent metal ion binding to all specific and unspecific binding sites on the D5 molecule as a whole. Consistent with an overlay of events of varying affinities, the K_D value from UV thermal melting is higher than the K_D value from fluorescence experiments. These results also identify and confirm the bulge region to be a site of somewhat higher metal ion binding affinity in D5. Yet, the experiments were conducted under aberrant buffer, pH as well as KCl conditions (Table 4). The $T_{m,K}$ values are shown to differ considerably under these varying conditions (Table 3). Melting temperatures not only depend on divalent, but also monovalent metal ion and buffer conditions. The dissociation constants are deduced from a series of melting temperatures T_m and therefore, depend on monovalent ion conditions in the background, too. The K_D values shown in Table 4 stem from experiments of different buffer conditions and background metal ion concentrations and thus, can not be directly and conclusively compared. Nevertheless, they can be taken as tendencies and allow plausible interpretations.

Table 4. Comparison of K_D values (mM) obtained from fluorescence measurements on D5-36-2AP and thermal melting of D5. Fluorescence experiments were performed in Standard Buffer (10 mM MOPS, pH 6.0, 10 mM KCl), whereas thermal melting studies were conducted in 100 mM KCl only.

	buffered	[KCl]	$K_{D,Mg}$	$K_{D,Ca}$
D5-36-2AP	+	10	0.20 ± 0.01	0.14 ± 0.01
D5	-	100	1.10 ± 0.21	1.18 ± 0.21

4. Single molecule FRET studies on the group II intron ribozyme *Sc*.ai5γ reveal a new paradigm for folding of large RNAs

4.1 Introduction

4.1.1 Folding of *Sc*.ai5γ

The splicing reaction and the secondary structure of group II introns and the folding of a *Sc*.ai5γ derivative (*Sc*.D135) have been introduced in detail in Chapter 1.7. In summary, the large *Sc*.D135 consists of five domains and folds to the native state in a slow and direct pathway that is devoid of kinetic traps[125] following an apparent two-step folding process involving an on-pathway intermediate (Figure 36).[130] The intermediate contains a collapsed form of D1, the formation of which is the rate-limiting step in folding.[132] Recently, it has been found that the initial and specific collapse of D1 is controlled by the κ-ζ region.[56] This small substructure lies at the center of D1, it is the docking site for D5 through tertiary contact formation and constitutes a part of the active site. Additionally, the κ-ζ region is also shown to be a metal ion binding pocket and prone to act as a metal ion dependent switch to initiate the cascade of folding events.[56, 225] Yet, the compaction rate constant is only slightly dependent on Mg^{2+} concentration, indicating the formation of an unstable folding intermediate that is captured upon binding of a Mg^{2+} ion to a weak metal ion binding site. *Sc*.D135 undergoes a fascinating folding pathway that mainly relies on the rate-limiting folding of a perfectly positioned subfragment to ensure proper active site formation for successful splicing already at the earliest stages of folding.

In its simplest form, the folding mechanism of the *Sc*.D135 ribozyme has so far been described by U ↔ I ↔ N (Figure 36b). The starting state U (unfolded) is defined as the extended form of the molecule in the presence of monovalent ions that adopts extensive secondary, but no tertiary structure. The rate-limiting and slow folding step U ↔ I, during which the large D1 undergoes massive compaction, dominates the entire folding pathway. The subsequent second step I ↔ N is known to be fast, but is not well characterized. So far, there is no evidence for any on- or off-pathway intermediate folding species next to and after I using standard biochemical techniques. It is generally believed that the remaining domains

fold rapidly into their native positions onto D1, which represents the folding scaffold. Folding of D2-D6 appears to take place in a concerted, but independent docking fashion.[132] This is somewhat surprising when considering the size and functionality of the isolated domains.

As mentioned previously, the molecule undergoes massive compaction upon folding. The

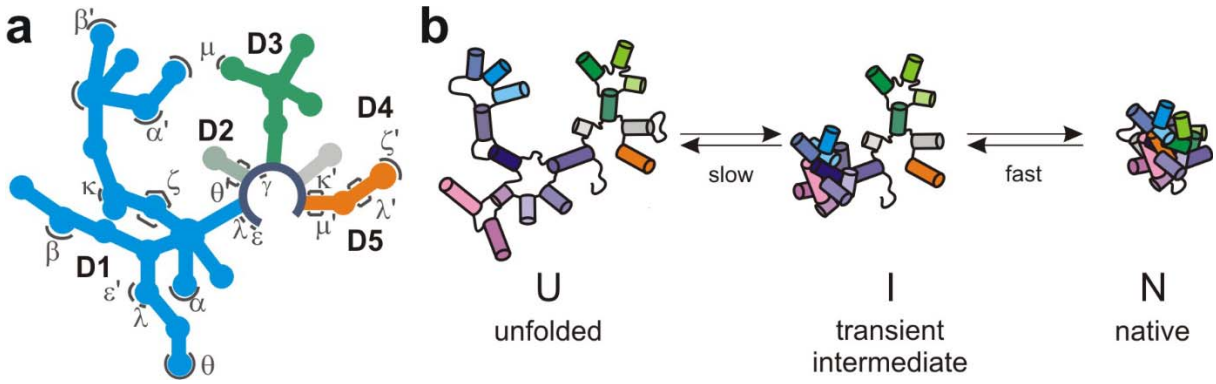


Figure 36. *Sc*.D135 secondary structure and folding pathway. **a** Depiction of the secondary structure of *Sc*.D135. The domains are colored in blue (D1), gray (D2);. green (D3), light gray (D4) and orange (D5) and the greek letters indicate positions of nucleotides involved in tertiary contacts. **b** Folding pathway of *Sc*.D135 at non-physiological conditions starting from the unfolded state (U) in the presence of monovalent ions only. Divalent metal ion addition initiates the slow folding step to the transient intermediate (I), in which D1 gets compacted, and the subsequent fast folding step to the native state (N). Figure b is adapted from [84].

hydrodynamic radius (R_H) of the molecule is reduced from 146 to 81 Å upon addition of monovalent metal ions and secondary structure formation (formation of U from a random coil structure). Addition of divalent metal ions induces the tertiary collapse and concomitantly, R_H is further minimized to 61 Å.[130] These folding events are accompanied by extensive conformational changes in the molecule over a distance range that is potentially accessible by FRET (Chapter 1.8.3, Figure 16). The FRET technique has been successfully applied for systems that are too large or too dynamic to be studied using other structural techniques.[226-228] These studies demonstrate the potential of FRET and imply possible applications of the technique to further and also larger RNA system, i.e. *Sc*.D135.

FRET requires the presence of fluorescent dyes. The intrinsic fluorescence of normal RNA bases is very weak and therefore, fluorescent groups need to be introduced into the RNA for luminescence studies.[228, 229] Modified fluorescent bases exist, but their luminescence intensity is usually still relatively weak and their fluorescence is likely to be quenched by base-pairing interactions.[230] Many fluorescence studies therefore rely on the site-specific introduction of very powerful and stable extrinsic fluorophores.[228] Dependent on the site-specific positioning of the fluorophore pair on the molecule, the distances and the fluctuations of selected parts in the molecules relative to each other can be measured by FRET.[139, 231] Such it becomes possible to observe the structural transitions during the biological function in real-time. This information is used to characterize the tertiary structure

of nucleic acids, to extract folding rates and to determine pathways and free energy landscapes of RNA.[232-234]

4.1.2 Techniques to label RNA with fluorophores

The nucleic acid base and backbone residues are usually of inert chemical nature and it thus becomes necessary to introduce extrinsic reactive groups into the molecules. There are a number of labeling strategies to attach fluorophores to nucleic acids.[235] For relatively short nucleotides (<40-100 nts) that are produced by solid-phase RNA synthesis it is possible to introduce fluorescent nucleotide analogs during synthesis.[236, 237] Alternatively, specific reactive groups (e.g. amine, thiol, thiophosphates) are incorporated into the RNA. After the chemical synthesis, these reactive groups can be coupled with a wide variety of chemical or spectroscopic probes. In this way it is possible to introduce terminal or internal fluorescent labels into base or backbone moieties.[231, 238] A large number of spectroscopic probes and modified phosphoramitides is commercially available from companies such as Molecular Probes and Glen Research. In addition, service laboratories such as Keck WM and Dharmacon offer customized chemical synthesis solutions for the most common applications.

The labeling strategies are more difficult in the case of large RNA (>100 nts) that can no longer be synthetically produced to satisfying yields and need to be isolated from cell extracts or originate from *in vitro* transcriptions.[239, 240] It is generally difficult - although partially possible[241, 242] - to incorporate a non-natural nucleic acid into a sequence during transcription, since the reaction is performed by RNA polymerase and this protein is carefully designed to only accept the natural bases as substrates and has an error rate of less than 1 per million.[243] In principle, it is possible to prime the transcription partially with guanosine 5'-monophosphorothioate (GMPS), although the transcription yield is usually reduced.[244, 245] The non-natural thiol group can then be reacted with thiol-reactive fluorophore derivatives. Alternatively, the RNA termini can posttranscriptionally be modified to phosphoamidates (5'-end)[235, 246] or oxidized to dialdehydes (3'-end)[126] and subsequently be labeled with fluorescence derivatives. These processes require multiple subsequent reaction steps that potentially reduce the RNA yields.

It remains a challenging task to site-specifically and covalently introduce a modification in the middle of an large RNA sequence.[235] In one approach, a long RNA can be divided into a set of smaller sequences, which are either *in vitro* transcribed or synthetically made and labeled using chemical synthesis. The RNA pieces are then linked in a template based ligation reaction using a DNA splint and T4 DNA ligase.[247] Yet, the efficiency of the ligation is

very construct dependent and this and other techniques can unfortunately not be used as general tools.[235]

A further technique makes use of the complementarity of bases to ultimately attach fluorophores to RNA rather than linking the fluorophores covalently to the RNA strand.[248] Short synthetic DNA-oligos carrying terminal fluorophores are designed with sequences that are complementary to specific single stranded loop structures within the long RNA. Given sufficient specificity of the DNA-oligo to an optimally chosen sequence in the large RNA, this approach allows efficient and site-specific labeling of RNA with fluorophores for biophysical studies through oligonucleotide hybridization.[248]

4.1.3 Single molecule FRET spectroscopy

Advantages of single molecule spectroscopy

The strong distance dependence of FRET has been utilized for biochemical studies in ensemble averaged solution to measure distances or overall distance changes and kinetics upon folding.[64] Recent advances in fluorescence microscopy[249] now also allow the detection of FRET at the single molecule level[250-252] and provide the tools to move a step forward towards the unraveling of the details of molecular processes. Using single molecule FRET (smFRET) spectroscopy, it becomes possible to record data from one single molecule isolated from its neighbors in real-time. The smFRET time traces reveal stochastic fluctuations between apparently distinct and preferred conformational states and allow the measurement of fluctuation rate constants under equilibrium conditions. In addition, smFRET experiments directly observe possible transient intermediates and parallel folding pathways including static and dynamic heterogeneity. This information is difficult or impossible to derive from bulk measurements since the contributions of short-lived and minor events are weak and become hidden when looking at the averaged signals of a large number of molecules at the same time.[148] Indeed, smFRET studies have revealed that identical molecules are able to follow different paths to new equilibrium states rather than that undergo a well-ordered set of events.[141, 144]

Total internal reflection fluorescence spectroscopy

There are two major techniques to achieve single molecule sensitivity: the confocal and the total internal reflection (TIR) fluorescence microscopy. The method of choice in this work is TIR fluorescence microscopy[148, 253] and hence, it will be summarized in more detail here. For information on confocal spectroscopy, please refer to [254].

A sample of FRET-labeled molecules is immobilized via the streptavidin-biotin interactions on a quartz slide and fixed in an inverted microscope (Figure 37). As an intense light source, a laser beam is focused on the sample through a prism at a particular angle of incidence (θ). This angle has to be larger than the critical angle (θ_c), beyond which the refracted beam does not penetrate the solution any more. The critical angle is determined by Snell's law[255, 256] using the indices of refraction of the slide n_s and the aqueous solution n_w:

$$\sin\theta_c = \frac{n_s}{n_w} \tag{5}$$

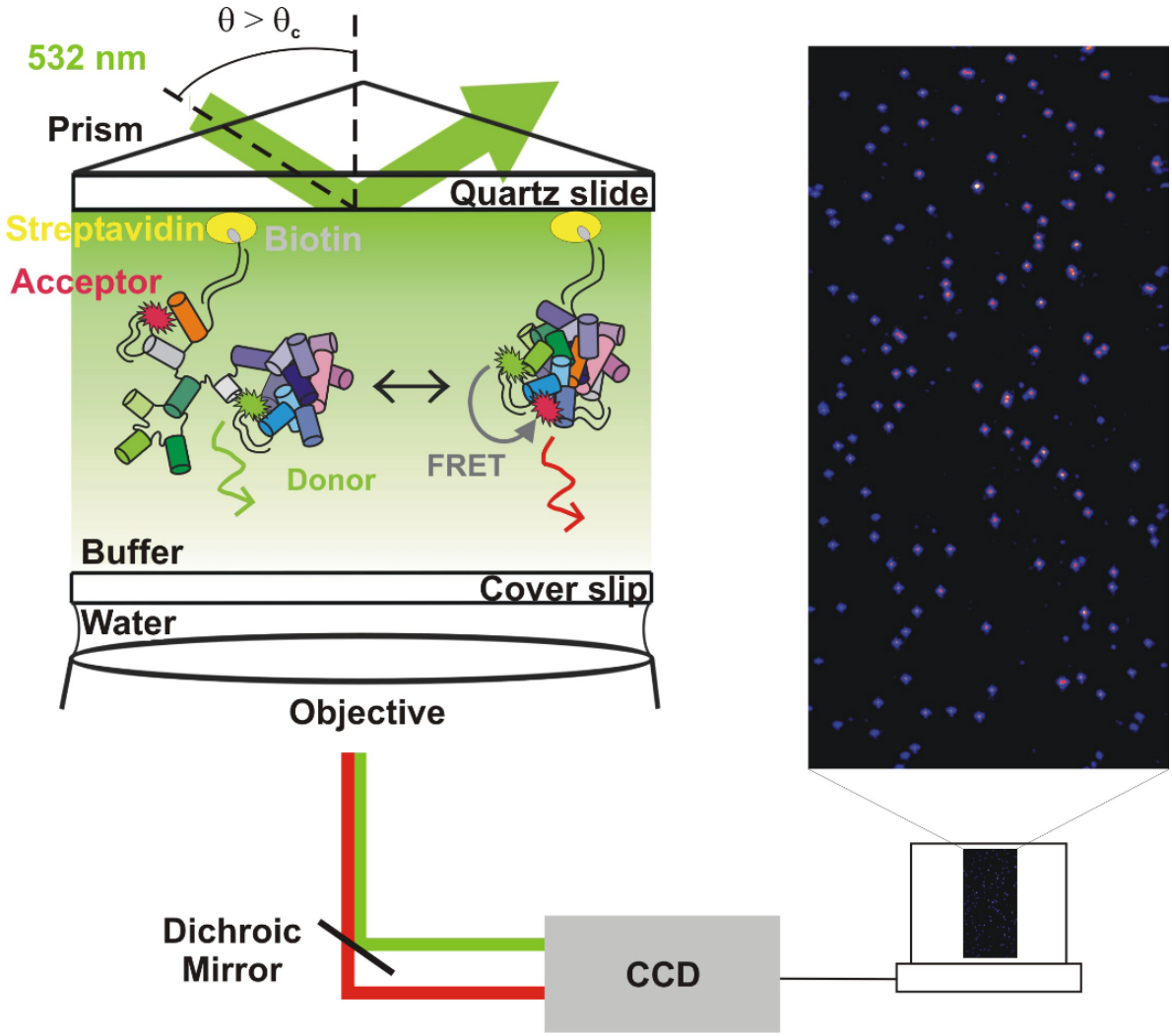


Figure 37. Single molecule detection by TIR fluorescence microscopy. The laser beam (incident angle $\theta>\theta_c$) is focused through the prism and reflected on the slide-water surface, generating an evanescent wave that excites immobilized molecules in an area up to 100 nm away from the slide surface. The emitted light is collected by the objective, split in donor and acceptor light using a dichroic mirror and recorded by the CCD camera. The molecules appear on the computer screen as bright dots on a black background. The section shown corresponds to the combined signals of the donor and acceptor emission.

The wave properties of the incident light generate an evanescent field at the slide/water interface that is typically able to excite molecules up to ~100 nm away from the slide surface.[257] The fluorescence light of these molecules is collected by the microscope objective onto an area detector (typically an intensified charged-coupled device (CCD) camera). Fluorescent signals from the donor and the acceptor molecules are separated with dichroic mirrors and detected as individual but congruent images on the CCD chip. The single molecules appear as tiny dots on the images that look like starlit skies and can be overlaid. The time-trajectories of relative smFRET efficiencies are calculated from integration of the donor and acceptor spots intensities and equation (6).

$$E = \frac{I_A}{I_A + \gamma I_D} \tag{6}$$

E describes the relative FRET efficiency, I_A and I_D are the acceptor and donor emission intensities, respectively, and γ is the acceptor-to-donor ratio of the product of fluorescence quantum yield and the instrument detection efficiency in the respective channel that typically is $\gamma \approx 1$.[258]

TIR fluorescence spectroscopy excites only a very thin sheet at the interface of the microscope slide and the solution (~100 nm), thereby reducing the background signal as well as the amount of photobleaching outside the observation volume. Furthermore, the TIR illumination and CCD camera imaging allow the observation of a wide field of single molecules without the need for sample scanning. Multiple molecules are detected in parallel and hence, the time necessary to collect a statistically significant number of single molecule time traces is relatively short. Next to all these advantages, TIR fluorescence spectroscopy is only limited by the time resolution of current CCD technology that is ~100 times lower compared to the fastest avalanche photodiodes (APDs) used in confocal fluorescence microscopy.[148]

4.1.4 Aim of chapter

The folding pathway and structure of the group II intron ribozyme *Sc*.ai5γ and its derivatives is being studied with a combined effort of various biochemical and structural techniques. The enormous size of this large ribozyme limits applications and interpretations of these standard techniques and studies are often conducted on isolated parts of the molecule only. Hence, there is a constantly growing amount of preliminary data, yet information on the entire molecule is, to a large extent, incomplete. Especially the overall folding pathway of this remarkable ribozyme structure is only partially understood.

Information on structure and folding of small ribozymes has to a significant amount been gained by FRET experiments and prove the feasibility of the FRET technique on RNA molecules. As the dimensions of *Sc*.D135 potentially match the distance range that is accessible by FRET, we set out to design and conduct a FRET study on this member of the group of large ribozymes that is only outsized by the largest ribozyme known to date, the ribosome. Since data is available on the initial and rate-limiting step in the folding pathway of *Sc*.D135, a stronger emphasis was set on elucidating the fast folding steps of this complex system.

4.2 Results

4.2.1 Finding the optimal FRET labeling scheme for smFRET experiments on folding of *Sc*.D135

It is crucial for any type of FRET experiment for successful structural and functional studies of nucleic acids to carefully choose a suitable labeling technique and the positions of the fluorophores on the molecule. The abovementioned oligonucleotide hybridization technique to efficiently label large RNAs with fluorophores[248] is a comparatively general and promising approach that could potentially be applied to our extremely large *Sc*.D135 system. A prerequisite are structurally and functionally unimportant hairpin loops on the RNA molecule that can be exchanged with modular hairpin loops of complementary sequence to the fluorophore-labeled DNA-oligos. A three-dimensional *Sc*.D135 model exists that combines all available biochemical data on *Sc*.D135 (at the time of publication in 2005) and is constructed from an extensive collection of known distance constraints.[191] Regions, for which no biochemical information on tertiary contacts is available, do not show up in this model. Several regions exist that have not been found to be involved in any type of importance for the ribozyme and can potentially be replaced by an unrelated nucleotide sequence without harm of function or structure.

Two hairpin loops in such regions in *Sc*.D135 were selected as sites for fluorescence labeled DNA-oligo attachment. For sufficient sequence specificity a hybrid duplex should have 12 or more basepairs.[248] The sequence length of the modular hairpin loops described in [248] (18 nucleotides) has proven to be long enough to complement a DNA oligonucleotide with sufficient stability for effective labeling and the DNA-RNA helix should be stably formed under all ionic conditions used in folding experiments. Therefore, the sequences indicated in [248] were adopted and introduced into the plasmid sequence of *Sc*.D135 at the appropriate positions by PCR cloning techniques using modified primers (*vide infra*).

4.2.2 *Sc*.D135-L1 and *Sc*.D135-L4

Sequence modifications in d2b of D1 and D4 are tolerated in catalysis

The two newly produced *Sc*.D135 derivatives are shown in Figure 38. One sequence contains the unrelated loop in a subdomain of D1 known as d2b and is named *Sc*.D135-L1 (L1: unrelated loop in D1, Figure 38b). The second sequence *Sc*.D135-L4 harbors the same unrelated loop sequence in D4 (Figure 38c), which is known to be a protein-binding site and

unimportant for self-splicing and whose sequence is already modified in *Sc*.D135 compared to the wild-type sequence in *Sc*.ai5γ (Section 1.7.2).

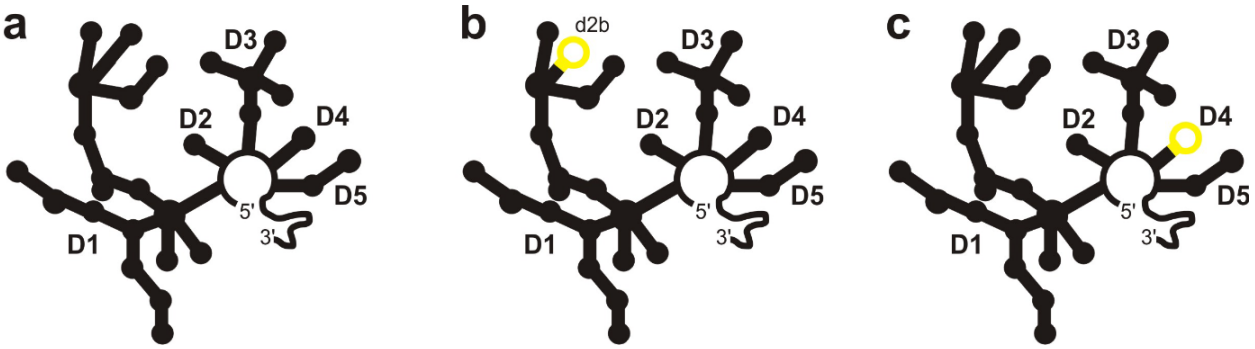


Figure 38. Schematic secondary structure depiction of **a** *Sc*.D135 **b** *Sc*.D135-L1 and **c** *Sc*.D135 showing the positions of the introduced hairpin loop (yellow) in subfragment d2b of D1 for *Sc*.D135-L1 and D4 for *Sc*.D135-L4.

The tertiary structure of ribozymes is conveniently reflected in the intrinsic cleavage activity of the molecules.[130] Only an active – and hence correctly formed – structure will result in product formation. The effect of the sequence change on the function of the ribozyme was tested for in chemical cleavage assays. The ribozyme *Sc*.D135 is designed to actively cleave an exonic substrate in a hydrolytic reaction and is able to perform experiments under single- and multiple-turnover conditions (STO and MTO, respectively).[129] In a *trans* cleavage assay, *Sc*.D135 cleaves the substrate 17/7 (24 nts) that contains the last 17 nts of the 5'-exon sequence (including both intron binding sites IBS1 and IBS2) and the first 7 nts of the intron, comprising the 5' splice site (Figure 39a).[129] The formation of product can be followed and quantified in an assay using radioactively ^{32}P-end-labeled substrate, PAGE analysis and subsequent phosphoimaging (Figure 39b).

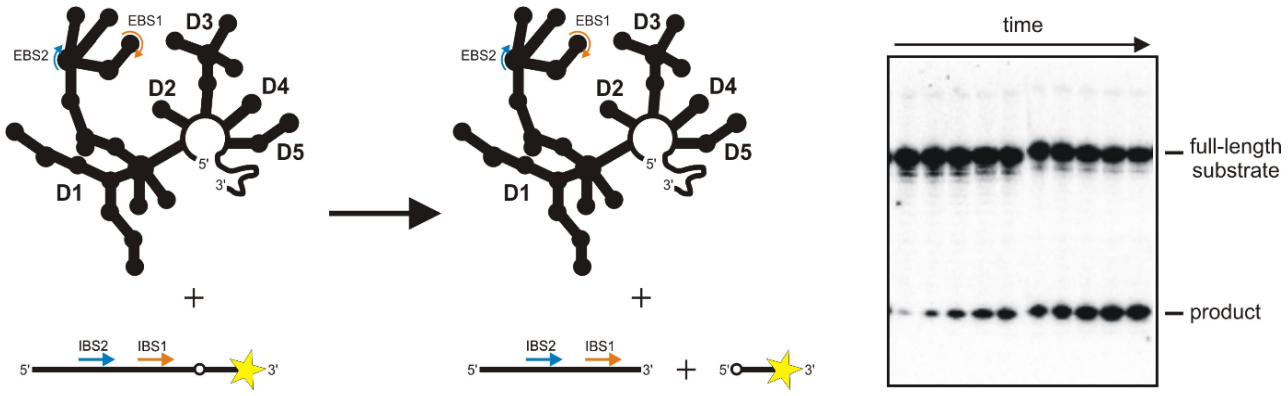


Figure 39. *Sc*.D135 cleavage reaction. **a** Schematic depiction of cleavage of 3'-^{32}P-labeled substrate 17/7 by *Sc*.D135. The radioactive label is depicted as a yellow star and the IBS1/EBS1 and IBS2/EBS2 sequences that form WC basepairs and serve for substrate recognition are depicted in orange and blue, respectively. The splice site on the substrate is indicated by a circle. **b** Section of a polyacrylamide gel showing the appearance of the product band in STO experiments with time. Slight shifts in band heights are due to a subtle misalignment of the combs that causes and determines the position of the wells during gel preparation.

The activity of *Sc*.D135-L1 and *Sc*.D135-L4 was recorded in STO experiments with either 3' or 5'-^{32}P-end labeled substrates. The kinetic parameters of cleavage and the amount of product formation were compared to *Sc*.D135 using the substrates 17/7 and a further substrate 17/7-dC, in which the C at the cleavage site is changed to a deoxy-C (dC). 17/7-dC is known to be cleaved by *Sc*.ai5γ derivatives at 5-15 times reduced rates[126] and is hence a slow-cleavable substrate. Both substrates (17/7 and 17/7-dC) were subjected to cleavage by *Sc*.D135, *Sc*.D135-L1 and *Sc*.D135-L4 and the results are summarized in Figure 40 and Table 5. *Sc*.D135-L4 shows similar rates to *Sc*.D135 ($k_{obs,Sc.D135\text{-}L4}$ = 0.54 ± 0.01 min^{-1} and $k_{obs,Sc.D135}$ = 0.53 ± 0.01 min^{-1}), whereas the sequence modification in d2b of D1 has a stronger, but still reasonably minor impact of 20% rate reduction on cleavage of 17/7 ($k_{obs,Sc.D135\text{-}L1}$ = 0.43 ± 0.04 min^{-1}). Cleavage rates are 15-20 times reduced and product formation drops from about 95% to 75% when introducing a deoxy-C at the splice site (using substrate 17/7-dC).

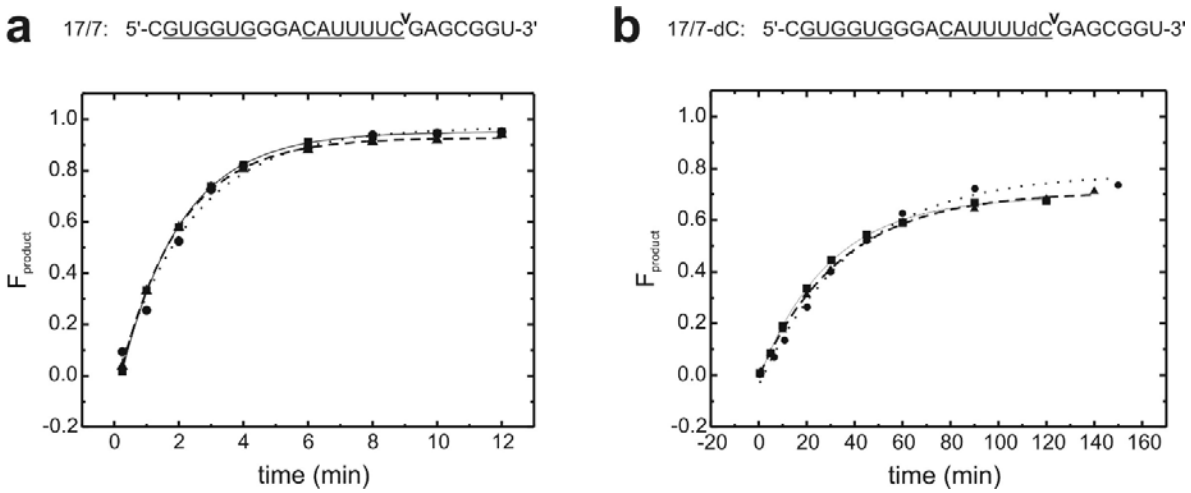


Figure 40. Product formation by *Sc*.D135 (●), *Sc*.D135-L1 (■) and *Sc*.D135-L4 (▲) for cleavage of substrates **a** 17/7, 5'-^{32}P labeled and **b** 17/7-dC, 3'-^{32}P labeled. The sequences of the substrates are shown on top of the graphs. The splice site is indicated with ^{v} and the underlined nucleotides represent IBS2 and IBS1 in this order. Single exponential fits are shown for *Sc*.D135 (solid line), *Sc*.D135-L1 (dotted line) and *Sc*.D135-L4 (dashed line). Cleavage reactions were performed in 80 mM MOPS pH 7.0, 500 mM KCl and 100 mM $MgCl_2$ at 42°C.

Table 5. Kinetic parameters for cleavage of 17/7 (5'-^{32}P-labeled) and 17/7-dC (3'-^{32}P-labeled).

substrate	ribozyme			k_{obs} (min^{-1})	% cleavage
	Sc.D135	*Sc*.D135-L1	*Sc*.D135-L4		
17/7	+	-	-	0.53 ± 0.01	95
	-	+	-	0.43 ± 0.04	97
	-	-	+	0.54 ± 0.01	93
17/7-dC	+	-	-	0.034 ± 0.001	71
	-	+	-	0.022 ± 0.001	78
	-	-	+	0.029 ± 0.001	72

Sc.D135-L1 and *Sc*.D135-L4 retain cleavage activity to a very large extent and the modifications in the sequence of d2b of D1 and in D4 do not greatly affect the overall function of the ribozymes. The loop d2b seems to be slightly more important than the stem-loop D4 structure, yet the d2b loop is still of relatively minor significance for the overall architecture of a functional ribozyme.

Fluorescent label on the substrate interferes with ribozyme cleavage

The ribozymes need to be labeled with a pair of fluorophores for any FRET experiments. The donor molecule Cy3 is covalently attached to a short oligonucleotide called Cy3-DNA, whose nucleotide sequence is complementary to the introduced loop sequence described above. Upon addition of Cy3-DNA to *Sc*.D135-L1 or *Sc*.D135-L4 and base pairing, the Cy3 label will non-covalently, but site-specifically be introduced in the ribozymes (Figure 41). In a first experiment, the acceptor molecule Cy5 was placed at the 5'-end of the substrates 17/7 and 17/7-dC.

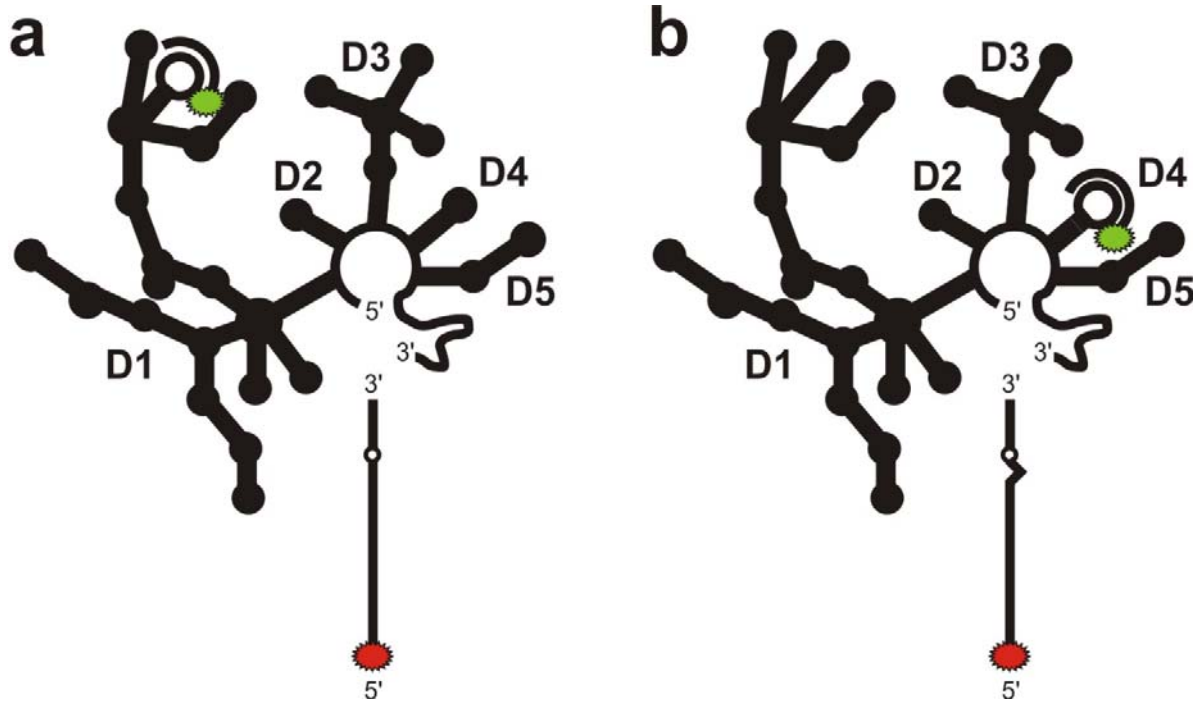


Figure 41. FRET labeling scheme shown on the secondary structure depiction of **a** *Sc*.D135-L1 with 17/7 and **b** *Sc*.D135-L4 with 17/7-dC. The Cy3-label (green) is attached to the ribozyme via base pairing of Cy3-DNA with the modified loop in D1 (a) or D4 (b). The Cy5-label (red) is placed on the substrate (17/7 in a or 17/7-dC in b). The splice site on the substrate appears as a circle and the deoxy-C substitution on 17/7-dC is indicated as an edge in the substrate.

Again, the introduction of the labels potentially interferes with cleavage activity. The performance of the Cy3-labeled ribozymes was tested and the cleavage of the Cy5-labeled substrates was recorded in cleavage assays analogous to the ones described above. The attachment of the Cy3-DNA to *Sc*.D135-L1 as well as *Sc*.D135-L4 do not impair cleavage of

the slow-cleavable substrate 17/7-dC (Table 6). In contrary, the cleavage rates increase by a factor of 1.5-2.5 and the product formations increase to 90-95%, indicating some beneficial stabilization by the double helix formation from the unstructured single stranded loop on the RNA by binding of the DNA-oligo. Unfortunately, cleavage of a Cy5-labeled substrate 17/7-Cy5 is severely blocked as shown both by the 2 fold reduced rates and, most strikingly, the amount of only 24 and 43% product formation! The 5'-positioning of the Cy5 label on the substrate apparently greatly interferes with substrate cleavage. Consistently, this is also noticeable when using the Cy5-labeled slow-cleavable substrate 17/7-dC-Cy5, where cleavage rates and product formation are, although maybe not as impressively, reduced (Table 6). The effect is hidden in the already reduced cleavage reaction due to deoxy-modification, whereas it becomes obvious when using the "wild-type" 17/7 sequence.

Table 6. Influence of fluorophores on cleavage reactions**.** All substrates are 3'-^{32}P-labeled.

substrate	ribozyme			Cy3-DNA	k_{obs} (min^{-1})	% cleavage
	*Sc.*D135	*Sc.*D135-L1	*Sc.*D135-L4			
17/7-dC	-	+	-	+	0.055 ± 0.003	95
	-	-	+	+	0.044 ± 0.005	91
17/7-dC-Cy5	+	-	-	-	0.033 ± 0.001	80
	-	+	-	-	0.025 ± 0.001	83
	-	-	+	-	0.035 ± 0.001	82
17/7-Cy5	-	+	-	-	0.18 ± 0.02	24
	-	-	+	-	0.25 ± 0.03	43

Both the stem-loops in d2b of D1 and D4 prove to be functionally unimportant sites and offer valuable positions to integrate unrelated nucleotide sequences. Fluorophore labeling of these loops via base pairing of fluorescently labeled DNA-oligos does not interfere with the catalytic reaction either. Yet, the substrate recognition is dominantly reduced when attaching a fluorophore label to the 5'-end of the substrate. Therefore, the placement of fluorophores on the substrate is not optimal. Additionally, FRET will only occur when the distance between acceptor to donor fluorophore is <100 Å, i.e. while the substrate is bound to the ribozyme. FRET will appear upon binding of the substrate and disappear upon substrate dissociation after cleavage. Such a set-up can potentially be used to measure binding, cleavage and/or dissociations rates as well as possible conformational changes upon cleavage, but not ribozyme folding - which is the aim of the project. These initial results were therefore used to develop an advanced FRET set-up that is better suited to study conformational changes upon folding of the ribozyme from the unfolded to the native state.

4.2.3 The construct *Sc*.D135-L14

Sc.D135-L14 sequence and labeling scheme

To be able to study folding of the *Sc*.D135 ribozyme it is necessary to place both the donor and acceptor fluorophore on the ribozyme itself. Two functional unimportant hairpin loops (d2b in D1 and loop in D4) that can be replaced by a modular hairpin loop were identified in experiments with *Sc*.D135-L1 and *Sc*.D135-L4 (Section 4.2.2). Therefore, the plasmid sequence of *Sc*.D135-L1 that already contains a modular hairpin loop in d2b of D1 was further modified to harbor a second unrelated hairpin loop in D4. This second modular

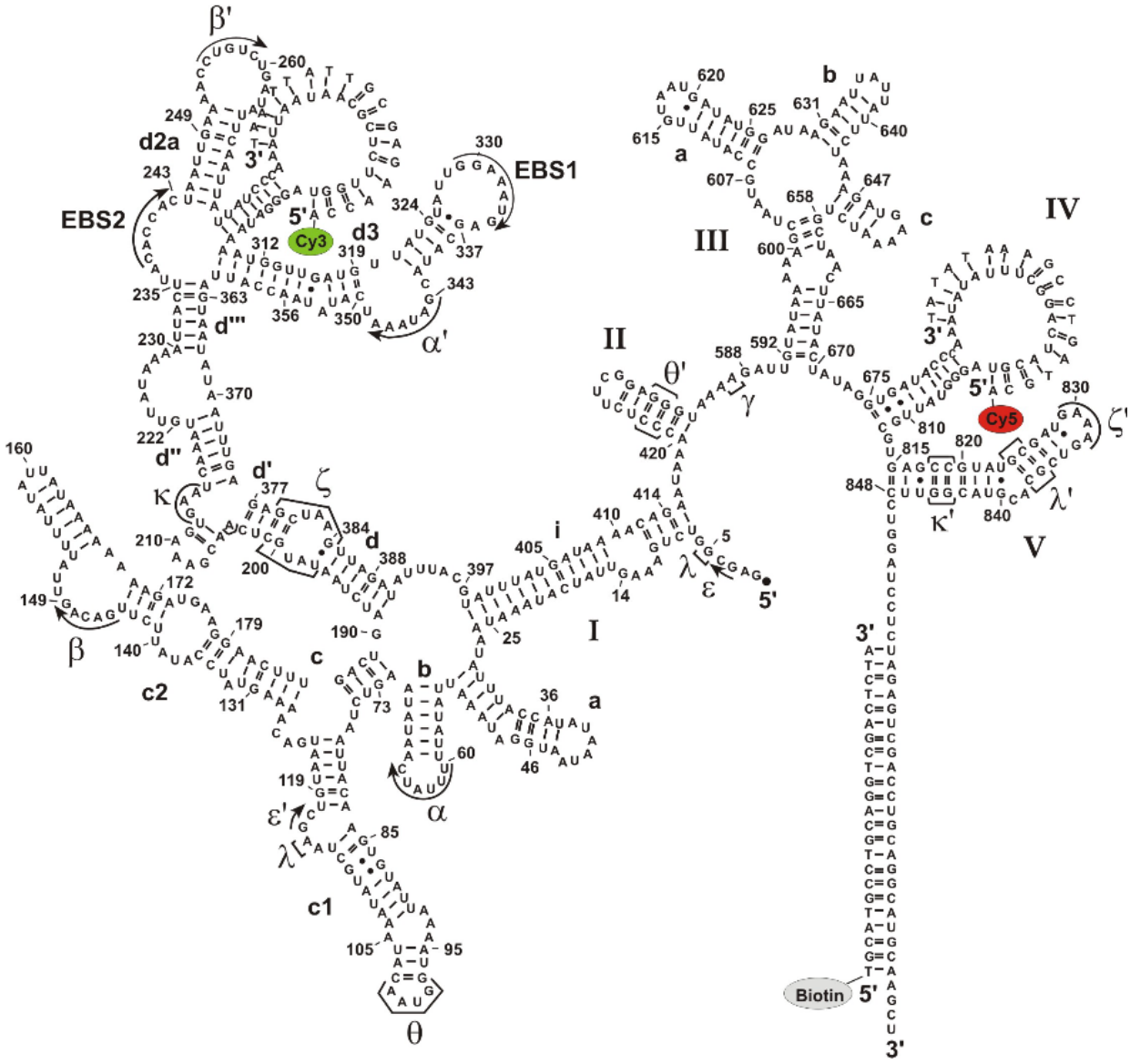


Figure 42. Secondary structure depiction of the *Sc*.D135-L14* ribozyme derived from the *Sc*.ai5γ group II intron in the *cox1* gene of *S. cerevisiae (Sc)*. Two modular loops are inserted in d2b of D1 and D4 and bind the DNA oligonucleotides with the fluorophores Cy3 (green) or Cy5 (red) at their 5'-ends. The numbering corresponds to the wild-type *Sc*.ai5γ sequence and the tertiary interactions are indicated in greek letters. The intron is elongated at its 3'-end and binds a 5'-biotinylated complementary DNA oligonucleotide (T-Biotin) for immobilization during smFRET experiments. The biotin label is indicated in gray.

loop is of the same sequence length as the first, but different in nucleotide sequence at 8 positions, where nucleobases were exchanged by their respective Watson-Crick partners.[248] A short oligonucleotide carrying the acceptor label and complementary to 18 nucleotides of this second modular loop was designed and called Cy5-DNA. The sequence differences in the first and second modular loop is necessary for sufficient and successful differentiation between Cy3-DNA and Cy5-DNA and sequence-specific binding of the oligos to their respective sites. The resulting construct is termed *Sc*.D135-L14 and the fully FRET-labeled ribozyme complex *Sc*.D135-L14* (*Sc*.D135-L14 with Cy3-DNA and Cy5-DNA), together with a further biotin-labeled oligo tether (T-Biotin) that is complementary to the 3'-overhang and necessary for immobilization in smFRET experiments (*vide infra*), is depicted in Figure 42.

Cy3-DNA and Cy5-DNA specifically and efficiently bind *Sc*.D135-L14

Native gels assays were performed to test for the sequence-specific binding of the fluorophore-labeled DNA oligos Cy3-DNA and Cy5-DNA to the internal positions on *Sc*.D135-L14. *Sc*.D135-L14 was prefolded in 100 mM $MgCl_2$ together with Cy3-DNA and/or Cy5-DNA and the complex formation was detected by shifts of fluorescence bands on a native FRET gel (Figure 43). No binding of Cy3-DNA or Cy5-DNA to *Sc*.D135 is detected in

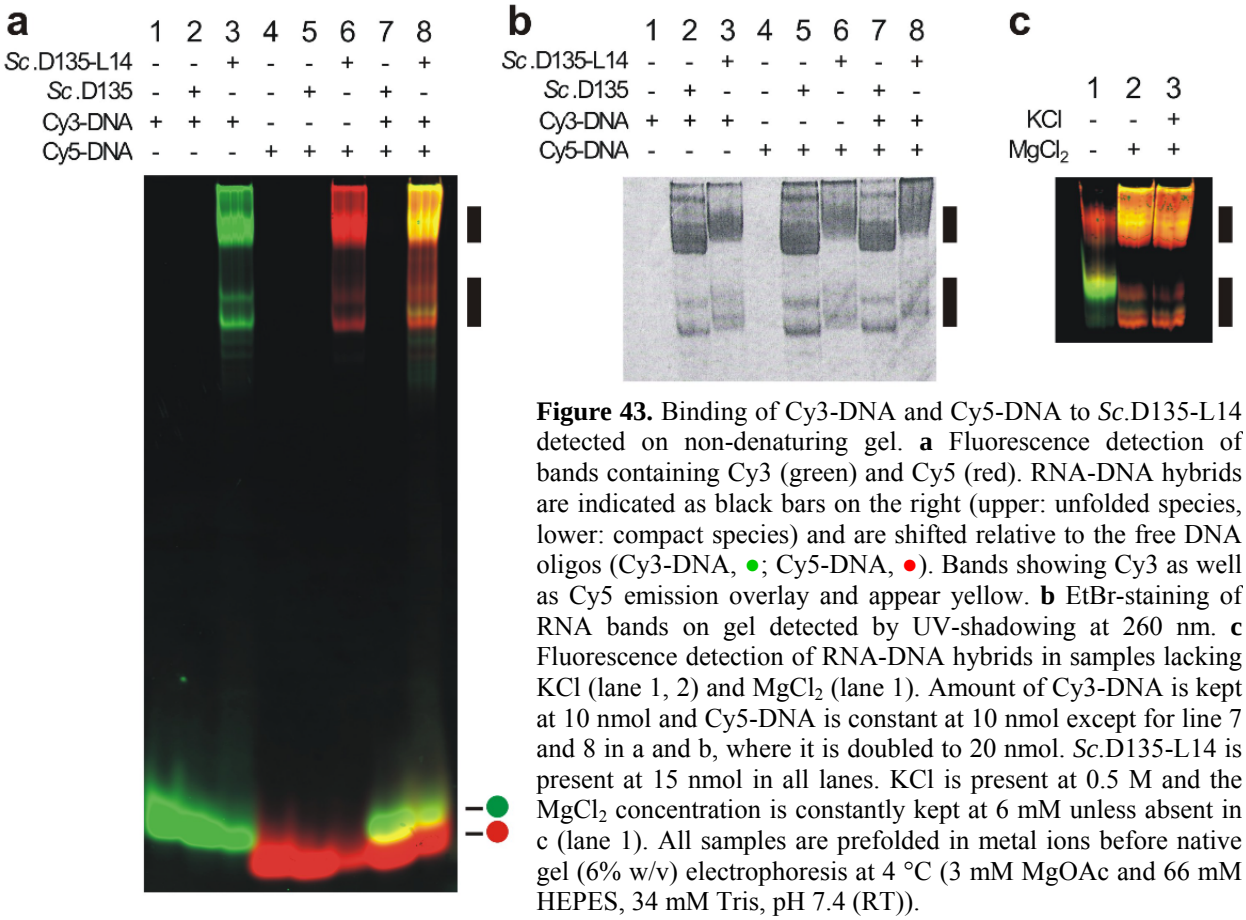


Figure 43. Binding of Cy3-DNA and Cy5-DNA to *Sc*.D135-L14 detected on non-denaturing gel. **a** Fluorescence detection of bands containing Cy3 (green) and Cy5 (red). RNA-DNA hybrids are indicated as black bars on the right (upper: unfolded species, lower: compact species) and are shifted relative to the free DNA oligos (Cy3-DNA, ●; Cy5-DNA, ●). Bands showing Cy3 as well as Cy5 emission overlay and appear yellow. **b** EtBr-staining of RNA bands on gel detected by UV-shadowing at 260 nm. **c** Fluorescence detection of RNA-DNA hybrids in samples lacking KCl (lane 1, 2) and $MgCl_2$ (lane 1). Amount of Cy3-DNA is kept at 10 nmol and Cy5-DNA is constant at 10 nmol except for line 7 and 8 in a and b, where it is doubled to 20 nmol. *Sc*.D135-L14 is present at 15 nmol in all lanes. KCl is present at 0.5 M and the $MgCl_2$ concentration is constantly kept at 6 mM unless absent in c (lane 1). All samples are prefolded in metal ions before native gel (6% w/v) electrophoresis at 4 °C (3 mM MgOAc and 66 mM HEPES, 34 mM Tris, pH 7.4 (RT)).

control samples (Figure 43a, lanes 2, 5 and 7) as fluorescence bands run at the same height as the free DNA-oligos (lanes 1, 4). On the other hand, the expected binding of Cy3-DNA and Cy5-DNA to *Sc*.D135-L14 is detected by shifts of fluorescence bands (lanes 3, 6 and 8), indicating the presence of several DNA-RNA complexes (*Sc*.D135-L14*) running at different heights in the gel and proving specific and very - although not quite 100% since there is still some unbound DNA-oligo present - efficient binding of DNA-oligos to *Sc*.D135-L14. The presence of doubly bound RNA-DNA hybrids is confirmed by the yellow bands in lane 8 originating from overlaid signals from Cy3 (green) and Cy5 (red).

RNA-DNA hybrid bands are also located using EtBr-staining of the native gel and UV-detection at 260 nm (Figure 43b). The free DNA-oligos are not detected by UV-shadowing because of reduced EtBr intercalation in the short single-stranded sequence length compared to the ds RNA-DNA hybrids and diffusion effect during EtBr-soaking and destaining. The longer sequence length of *Sc*.D135-L14 and binding of DNA-oligos to *Sc*.D135-L14 is confirmed by the general shift of bands compared to *Sc*.D135 (Figure 43b, lanes 2 vs. 3, 5 vs. 6 and 7 vs. 8), indicating RNA-DNA hybrids of bigger size. *Sc*.D135-L14 shows bands that are slightly more diffuse and one RNA-only band (second from top) completely disappears compared to *Sc*.D135. This is due to sequence modification from *Sc*.D135 to *Sc*.D135-L14 (data not shown), but fluorescence detection confirms DNA-oligo bound *Sc*.D135-L14 and derivations depicted in Figure 43b, lanes 2 vs.3 and 5 vs.6, are due to the combination of sequence modification and DNA-oligo binding. The double oligo bound RNA in Figure 43b, lane 8 is shifted another size gap compared to either singly Cy3-DNA or Cy5-DNA-bound RNA in lanes 3 and 6, respectively, also clearly confirming DNA-oligo binding via UV-shadowing.

A minor amount of RNA or RNA/DNA hybrids consistently sticks to the wells and does not move into the gel (Figure 43, top bands). This band might indicate the presence of potential aggregates generated during sample preparation or limited mobility of the large RNA/DNA hybrids even in low percentage native gels. Most probably, this static band is linked to the high amount of metal ions in the samples or not perfectly optimized buffer conditions.

Mg^{2+} concentration dependent changes in Cy3- and Cy5-DNA bound *Sc*.D135-L14 conformations could not be resolved and are found to be difficult to detect in the present native gel assay (data not shown). Nevertheless, distinct band shifts are detected upon addition of metal ion (Figure 43c, lanes 1 vs. 2 and 3), indicating large conformational differences between random coil and higher order structures formed by the ribozyme. As

previously suggested, it is indeed the presence of metal ions that causes the DNA-oligo bound *Sc*.D135-L14 molecules showing no mobility in the native gel assay and appearing as the static bands mentioned above.

The fluorophore-labeled *Sc*.D135-L14* is catalytically competent

Analogous to STO cleavage experiments with earlier constructs, the impact of the sequence modifications and binding to fluorophore and biotin-labeled DNA-oligos was measured in activity assays using radioactively labeled substrates. Depending on the position of the radioactive ^{32}P label on the substrate, some differences in cleavage efficiency were detected (Figure 44, Table 7).

When using 3'-^{32}P-labeled 17/7, the cleavage rate k_{obs} drops 17% upon sequence modification of *Sc*.D135 (k_{obs} = 0.41 ± 0.02 min^{-1} and k_{obs} = 0.37 ± 0.02 min^{-1} for *Sc*.D135 and *Sc*.D135-L14, respectively). The cleavage rate decreases a further 56% (63% compared to *Sc*.D135) when all DNA-oligos (Cy3-DNA, Cy5-DNA, T-Biotin) are bound in *Sc*.D135-L14*. Simultaneously, the fraction of product formed decreases from 70% to 66% to 55%. This indicates some inhibition by the sequence modification or the DNA-oligonucleotide, but could possibly also be caused by the additional ^{32}P-labeled C that is attached to the 3'-end of 17/7. Yet C is the next nucleotide following in the wild type *Sc*.ai5γ sequence and the slight change on the phosphate should not have such a measurable effect. However, considering the distance of 7 nts from the 3'-end-label to the splice site, it would maybe be more advantageous to place the label on the 5'-end.

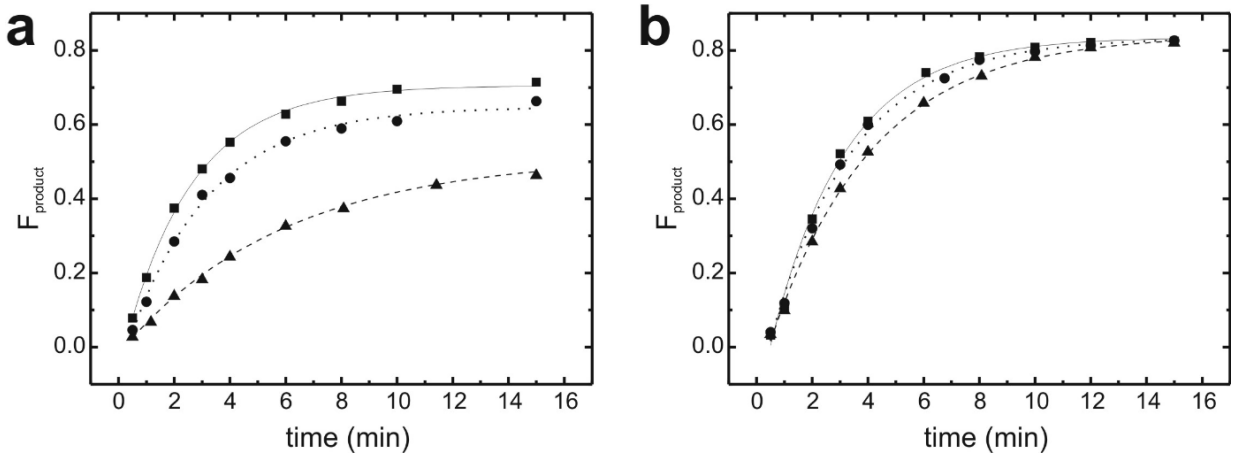


Figure 44. Cleavage of 17/7 by *Sc*.D135-L14. **a** Cleavage of 3'-^{32}P-labeled 17/7 by *Sc*.D135 (■, solid line), *Sc*.D135-L14 (●, dotted line) and fully labeled *Sc*.D135-L14* (▲, dashed line) **b** Cleavage of 5'-^{32}P-labeled 17/7 by *Sc*.D135 (■, solid line), Sc.D135-L14 (●, dotted line) and fully complexed *Sc*.D135-L14* (▲, dashed line) Lines depict single exponential fits and reactions were performed in 80 mM MOPS, pH 7.0, 500 mM KCl, 100 mM $MgCl_2$ and at 42°C.

Indeed, in cleavage experiments with 5'-^{32}P-labeled 17/7, the cleavage rates are k_{obs} = 0.40 ± 0.02 min^{-1}, k_{obs} = 0.37 ± 0.02 min^{-1} and k_{obs} = 0.27 ± 0.01 min^{-1} for *Sc*.D135, *Sc*.D135-L14 and for *Sc*.D135-L14*, respectively and only decrease 7.5% and 27% (32.5% compared to *Sc*.D135) upon sequence modification and binding of fluorophore and biotin labeled DNA-oligos, respectively (Table 7). Nearly no change in product formation is detected when using 5'-^{32}P labeled 17/7 as product formation is consistently measured to be 83-84%.

Cleavage reactions with the slow cleavable substrate were only carried out with 3'-^{32}P-labeled 17/7-dC and result in k_{obs} = 0.05 ± 0.01 min^{-1} and 87% product formation by *Sc*.D135-L14* (Table 7) after ~80 min reaction time (see Figure 40b, Table 6 and 7 for comparison with other *Sc*.D135 constructs).

Table 7. Cleavage rates and product formation by *Sc*.D135, *Sc*.D135-L14 and *Sc*.D135-L14*.

substrate	*Sc*.D135	*Sc*.D135-L14	*Sc*.D135-L14*	k_{obs} (min^{-1})	%$_{cleavage}$
3'-^{32}P-17/7	+	-	-	0.41 ± 0.02	70
	-	+	-	0.34 ± 0.02	66
	-	-	+	0.15 ± 0.01	55
5'-^{32}P-17/7	+	-	-	0.4 ± 0.02	84
	-	+	-	0.37 ± 0.02	84
	-	-	+	0.27 ± 0.01	83
3'-^{32}P-17/7-dC	-	-	+	0.05 ± 0.01	87

The *Sc*.D135-L14* ribozyme requires Mg^{2+} for cleavage

Group II introns are known to require high ionic strength for optimal activity *in vitro*. This is thought to be due to a high Mg^{2+} requirement for compaction of D1[44, 94] and additionally, stabilization of the native state.[132] The Mg^{2+} dependency of k_{obs} for *Sc*.D135-L14* was measured in the range from 0 to 100 mM Mg^{2+} in STO cleavage experiments, both for 3' and 5'-^{32}P labeled 17/7 (Figure 45, Table 8). *Sc*.D135-L14* that was prefolded in Mg^{2+} shows dissociation constants K_{Mg} = 39.6 ± 2.6 mM and K_{Mg} = 52.7.6 ± 3.4 mM for Mg^{2+} dependent cleavage of 3'- or 5'-^{32}P labeled 17/7, respectively. Again, 5'-^{32}P labeled 17/7 substrate was cleaved faster and to a higher extent than 3'-^{32}P labeled 17/7, although $k_{obs,max}$ for 5'-^{32}P labeled 17/7 was reduced in the set of titration experiments as compared to the experiments mentioned above (Figure 45, Table 8). The buffer used in this set of experiments was found to contain a slightly reduced pH (0.3 pH units), which most probably is also the reason for the higher K_{Mg} obtained in this assay compared to the one with 3'-^{32}P labeled 17/7 (Table 8).

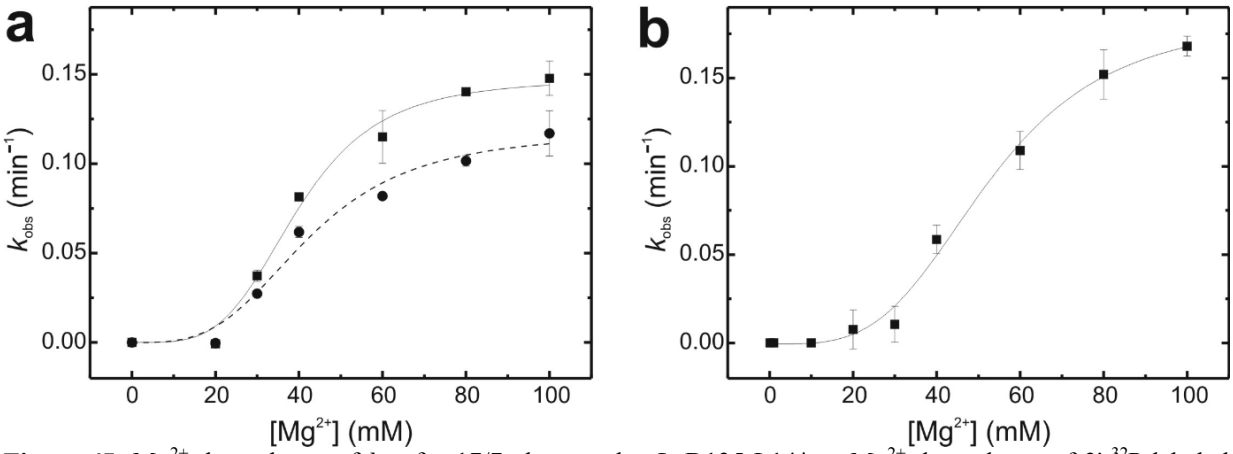


Figure 45. Mg^{2+} dependency of k_{obs} for 17/7 cleavage by *Sc*.D135-L14*. **a** Mg^{2+} dependence of 3'-^{32}P-labeled 17/7 cleavage by *Sc*.D135-L14* preincubated in Mg^{2+} (■, solid line) or if Mg^{2+} is added concomitantly with substrate (●, dashed line) **b** Mg^{2+} dependence of 5'-^{32}P-labeled 17/7 cleavage preincubated in Mg^{2+} (■, solid line). Lines depict Hill fits and reactions were performed in 80 mM MOPS, pH 7.0, 500 mM KCl and at 42°C.

If *Sc*.D135-L14* is not prefolded in Mg^{2+}, then the molecules need to fold first before cleavage can occur during the time course of cleavage acquisition. Cleavage rates would potentially be reduced since the rate-limiting step is folding of the molecules. If Mg^{2+} is added concomitantly with 17/7 at the starting point of the reaction, the fit parameters slightly change to $k_{obs,max}$ = 0.12 ± 0.01 min^{-1} and K_{Mg} = 42.5 ± 5.4 mM for 3'-^{32}P labeled 17/7 (Figure 45a, Table 8). These values are largely within the error limits as found for prefolded *Sc*.D135-L14* and do not indicate a need of any extra metal ions for catalysis that are not already used for folding. Yet, the detection of a potentially small amount of a few metal ions specifically needed for catalysis compared to folding is a difficult task for a very large RNA system and especially, in a metal ion background as high as 100 mM. It is not surprising that detection of metal ions for cleavage only is not easily possible by the relatively simple cleavage assay described here and metal ions specifically needed for catalysis can not be ruled out.

Table 8. Mg^{2+} dependencies of *Sc*.D135-L14* cleavage.

substrate	prefolding	$k_{obs,max}$ (min^{-1})	$\%_{cl,\ max}$	K_{Mg} (mM)
3'-^{32}P-17/7	+	0.15 ± 0.01	55	39.6 ± 2.6
	-	0.12 ± 0.01	54	42.5 5.4
5'-^{32}P-17/7	+	0.18 ± 0.01	85	52.7 ± 3.4

Fluorophores attached to DNA-oligos and bound to *Sc*.D135-L14* show typical fluorescence anisotropy values

FRET is dependent on three individual parameters: the spectral overlap of donor emission and acceptor excitation spectra, the distance between donor and acceptor, as well as the relative orientation of the donor and the acceptor dipole moments. The κ^2 value is an orientation factor describing the spatial relationship between the fluorophore dipoles in three-dimensional space. The fluorophores are usually assumed to be freely rotating and not experiencing any conformational constraint that would fix their dipoles in a certain orientation. Then, an average value of $\kappa^2 = 2/3$ can be applied to calculate FRET distances as described in Section 1.8.3. Variations in κ^2 within 1-4 result in an uncertainty in the calculated distance of no more than 35 %. Severe deviations only result if the dipoles are oriented perpendicular to each other, i.e. $\kappa^2 = 0$.[134]

Measurements of fluorescence anisotropies of the donor and acceptor indicate the rotational freedom of the fluorophores and can thereby set limits on κ^2 and minimize uncertainties in the calculated distance. The fluorophore is excited with vertically polarized light and both the vertical (vv) and the horizontal (vh) emission polarizations are captured and used to calculate anisotropy values r (eq.12, Section 5.5.5). If the fluorophores are rotating freely, vv and vh will become equal and the r value will be close to 0 (depolarized). If the fluorophores experience a rotational constraint, this will be reflected in an r value close to the maximal theoretical value of 0.4 (polarized).

The r values for both Cy3-DNA and Cy5-DNA are 0.27 ± 0.01 (Table 9). Upon binding to *Sc*.D135-L14, polarization of Cy5-DNA does not change while Cy3-DNA experiences some more polarization (r = 0.33 ± 0.02).

Table 9. Anisotropy values (r) for DNA-oligos in the absence or presence of *Sc*.D135-L14

	Sc.D135-L14	r
Cy3-DNA	-	0.27 ± 0.01
	+	0.33 ± 0.02
Cy5-DNA	-	0.27 ± 0.01
	+	0.27 ± 0.01

The rotational freedom of a dye is determined by the rotation of the fluorophore itself, as well as the overall rotation of the molecule, to which the fluorophore is attached. The dyes become constrained when covalently bound to the short DNA-oligos due to the increased size of the molecule compared to free fluorophores and thus, the decreased rate of rotational diffusion during the lifetime of the excited state. Additional hydrogen bonding of the DNA-oligos to *Sc*.D135-L14 introduces no further polarization for Cy5, whereas the dipole of Cy3 becomes slightly more fixed, indicating 20% less rotational freedom of the Cy3 dye within the compact network of D1 compared to the position of Cy5 on D4 that is located on the surface of the molecule. The *Sc*.D135-L14 system described here shows typical anisotropy

values for fluorophores coupled to nucleic acids.[244] It is important to note that the fluorophores attached to the DNA oligonucleotides do experience no or only moderate additional constraints by binding to the enormous *Sc*.D135-L14 RNA network. Therefore, the Sc.D135-L14* complex behaves comparable to smaller RNA systems and no delay in rotation caused by the huge size of the complex can be detected. At the same time, changes in anisotropy directly confirm binding of Cy3-DNA to *Sc*.D135-L14. In any case, a feasible application of Sc.D135-L14* in FRET and FRET distance calculations is confirmed.

The construct *Sc*.D135-L14* shows a Mg^{2+} dependency for folding

The folding behavior of freely diffusing *Sc*.D135-L14* in solution was characterized at 42 °C to determine the role of Mg^{2+} independently of catalysis. The fluorescence emission spectrum of *Sc*.D135-L14* in the absence of Mg^{2+} shows two bands that correspond to the donor and the acceptor fluorophores, respectively (Figure 46a). The donor intensity at 565 nm is higher than the acceptor at 665 nm. Upon addition of 100 mM Mg^{2+} the donor intensity decreases slightly while the acceptor intensity increases. As a result, the observed FRET ratio increases exponentially from 0.275 to 0.291 with a rate constant of k_{obs} = 0.9 ± 0.1 min^{-1} (Figure 46b). This rate is in very good agreement with the rate limiting step of D1 compaction obtained from hydroxyl radical footprinting, NAIM assays and ribozyme activity assays.[84, 125, 132]

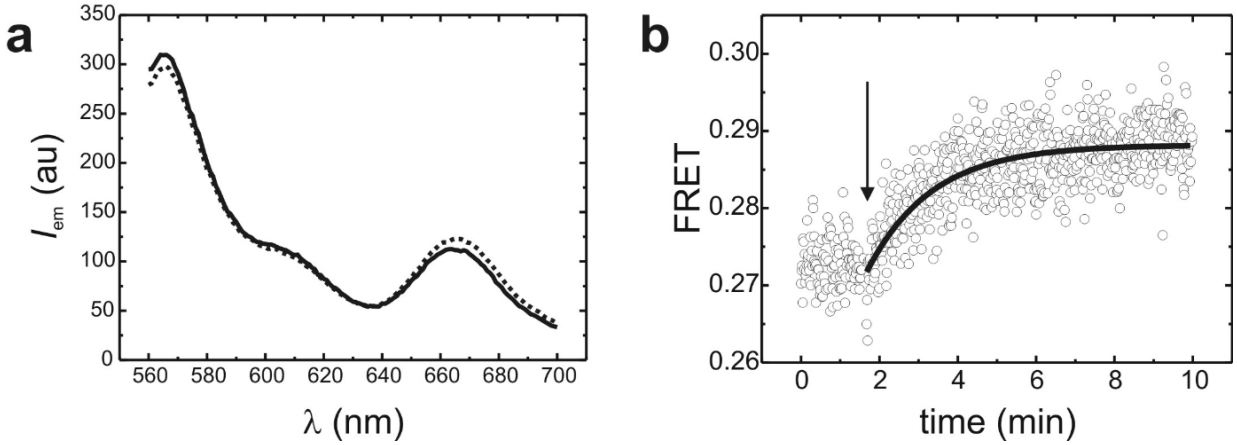


Figure 46. Bulk fluorescence measurements of *Sc*.D135-L14*. **a** Emission spectra (λ_{ex} = 555 nm) of *Sc*.D135-L14* in the absence (0 mM, solid line) and the presence of $MgCl_2$ (100 mM, dashed line) show correlating shifts in maxima at 565 nm (Cy3) and 665 (Cy5) caused by FRET. **b** FRET change upon addition of 100 mM Mg^{2+}. The arrow indicates the time point of addition of Mg^{2+}. The single exponential fit of the experimental data is shown as a black line. Data is recorded in 80 mM MOPS, pH 7.0, 500 mM KCl, $MgCl_2$ as indicated and at 42 °C.

The FRET increase to 0.291 in this bulk experiment suggests that a conformational change takes place in the presence of Mg^{2+} ions during which the relative distance between the fluorophores on the RNA decreases. To determine the Mg^{2+} dissociation constant of this distance change, the observed FRET ratio is plotted as a function of Mg^{2+} concentration

(Figure 47a) and fit with the Hill equation (eq. 11, Section 5.5) yielding K_{Mg} = 42.3 ± 1.3 mM. This value is again in very good agreement with the K_{Mg} from the cleavage reaction described above (Table 8) as well as previously reported values for *Sc.*D135.[125, 156]

The folding rates k_{obs}, calculated by assuming single exponential kinetics of FRET increase upon Mg^{2+} addition (as depicted in Figure 46b), show a decreasing tendency with increasing Mg^{2+} concentration (Figure 47b). The rates at Mg^{2+} concentrations <20-30 mM most probably do not describe folding but are rather of convectional or concentration and temperature related nature as the absolute FRET changes at these low Mg^{2+} concentrations are minimal (Figure 47a) and errors for these rates are big (Figure 47b). Therefore, this rate analysis fails to identify folding rates below 20-30 mM but reliably serves to state a roughly constant rate of folding k_{obs} = 0.9 ± 0.1 min^{-1} at 50 mM $MgCl_2$ and above.

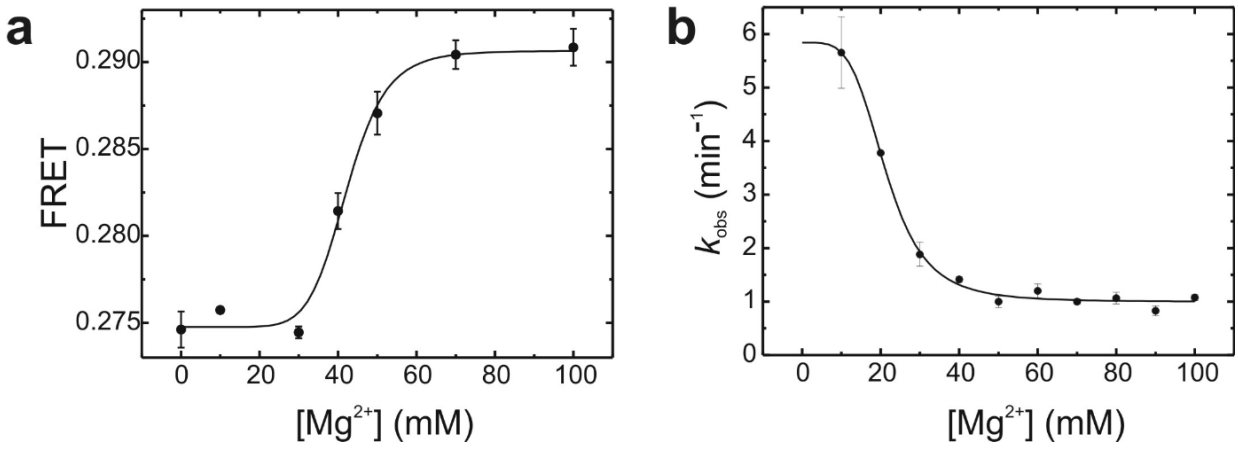


Figure 47. Mg^{2+} dependencies of FRET change and folding rates k_{obs}. **a** Maximal change in FRET upon increasing amounts of Mg^{2+}. The line depicts fit with the Hill equation giving K_{Mg} = 42.3 ± 1.3 mM. **b** Changes in FRET upon Mg^{2+} addition show decreasing rates. The fit with the Hill equation is depicted as a solid black line and results in K_{Mg} = 21.4 ± 0.6 mM. Samples are measured in 80 mM MOPS, pH 7.0, 500 mM KCl, $MgCl_2$ as indicated and at 42 °C.

Fluorophores in folded *Sc.*D135-L14* keep an average minimal distance of 45-50 Å

Steady-state FRET experiments allow the detection of qualitative structural changes and characterization of the overall RNA folding kinetics. Quantitative information about the absolute distances between the fluorophores on the other hand is derived from a different assay using the time-resolved FRET (trFRET) technique based on the measurement of donor fluorescence lifetimes.[148] First, the time-resolved donor fluorescence decay is measured in a sample containing donor-only labeled ribozyme. Second, the fluorescence decay of the donor is measured in a donor-acceptor labeled sample under identical conditions. Due to FRET from donor to acceptor, which is dependent on the donor-acceptor distance, the donor fluorescence lifetime in the doubly labeled sample is shorter and the decay will be faster (Figure 48a). Motions of large size molecules are slow compared to the time scale of a typical

donor fluorophore lifetime in the excited state (nanoseconds). Therefore, the donor-acceptor distance appears fixed and will be characterized by a unique donor fluorescence lifetime. If there are multiple distinct conformational isomers in equilibrium, then every conformer will contribute its specific donor lifetime distribution to the overall fluorescence decay. These individual contributions can be mathematically deduced assuming gaussian distributions and result in mean distance values and relative abundance of the isomers - as long as the individual donor lifetimes (and hence donor-acceptor distances) do not overlap to an inseparable extent.

The distance distribution of *Sc*.D135-L14* deduced from trFRET experiments appears as a single and very broad band. The distance distribution band can be fitted assuming a single or double gaussian distribution to approximately equally satisfactory extent. The simplest interpretation is the assumption of the presence of one conformational isomer that shows a broad bandwidth in distance. The mean distance of this single conformer changes from 136 Å to 94 ± 4 Å upon addition of 1 mM Mg^{2+}, indicating a large collapse upon divalent metal ion addition (Figure 48b). Yet, these calculated values are out of the range that FRET is most sensitive (~20-80 Å, Figure 16c) and can not be directly adopted. They basically only indicate an extended conformation of *Sc*.D135-L14* at low Mg^{2+} concentrations. Further Mg^{2+} addition up to 100 mM decreases the mean distance to a minimal value of 45 ± 12 Å and the distance change is described by a Mg^{2+} dependency resulting in K_{Mg} = 37 ± 12 mM.

These distance values nicely follow the hydrodynamic radius data on *Sc*.D135[130] as well as match the Mg^{2+} dependencies in cleavage reactions (Table 8) and described in the literature.[125, 156] The large errors on all these values clearly indicate a considerable uncertainty of the trFRET parameters. As already mentioned, it is equally possible to state a

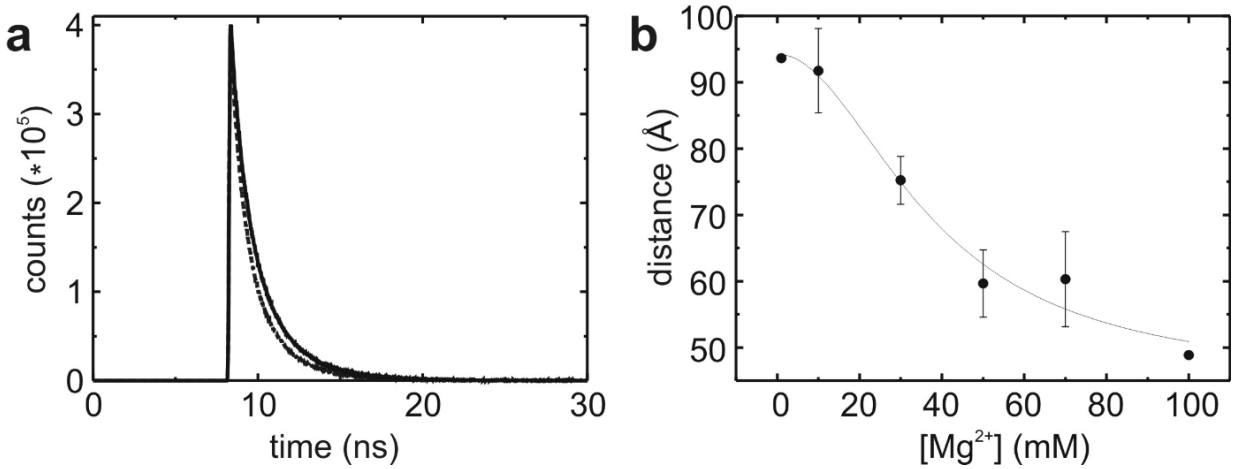


Figure 48. Time-resolved FRET and averaged distance data on *Sc*.D135-L14*. **a** Donor lifetime decay data in the absence (solid line) and the presence (dashed line) of acceptor. **b** Mean distance reduction upon Mg^{2+} addition assuming the presence of one single conformer. The error bars stem from two independent measurements. The solid line depicts a Hill fit and results in K_{Mg} = 37 ± 12 mM. Experiments were recorded in a buffer of 80 mM MOPS, pH 7.0, 500 mM KCl at 25 °C.

mixture of two isomers of similar and hence, overlapping mean distances and the interpretation is not straightforward. As trFRET experiments measure averaged signals from a lot of molecules and potential minor populations of extended or more compact conformations can not be ruled out, the measured value of 45 Å at 100 mM is an average and it is possible that the actual distance of the fluorophores on the main conformation of *Sc*.D135-L14* is actually shorter than 45 Å.

Taken together, all the above results demonstrate that the *Sc*.D135-L14* behaves like the *Sc*.D135 ribozyme regarding catalysis, folding, and Mg^{2+} requirements.[125, 156] We thus proceeded to study the conformations, their dynamics and the folding pathway of the *Sc*.D135-L14* ribozyme by single molecule fluorescence.

4.2.4 Immobilization and smFRET experimental procedure

To observe localized fluorescence signals of single molecules over an extended period in real-time, the molecules need to be immobilized on a surface. The biotin-streptavidin interaction has a K_D in the order of $4*10^{-14}$ M and is the strongest non-covalent biological interaction known.[259] The extraordinarily strong affinity of this interaction is used to immobilize ribozymes on the glass surface and has successful been exploited in TIR fluorescence microscopy before. The ribozyme sequence is elongated by a 37 nt extension at the 3'-end, to which a complementary DNA-oligo tether carrying a biotin moiety at its 5'-end was designed (Figure 42). This 24 nt long sequences called T-Biotin is attached to the ribozymes by the same DNA-oligo hybridization technique that has already been used for binding of fluorophores to site-specific positions on the RNA. The glass slide is coated with a layer of streptavidin prior to smFRET experiment and the *Sc*.D135-L14* molecules will therefore be attached and immobilized when applied to the glass slide.

Sc.D135-L14* ribozymes were incubated in the presence of specific amounts of Mg^{2+} in the range of 1-100 mM at 42°C to induce folding prior to immobilization and smFRET experiments. The sample was then applied to the streptavidin coated slide at room temperature (RT) and the molecules became immobilized on the quartz surface via the biotin-streptavidin interaction. All RNA system studied today exhibit only negligible perturbation from surface immobilization.[260] The optimal surface density of the molecules on the slide for optical resolution was determined by adjustments of concentrations in ribozyme samples to 50-100 pM. The immobilized molecules on the slide were washed into a solution containing an oxygen scavenger system to reduce photobleaching effects and fluorescence

signal acquisition was performed upon laser excitation of selected spots on the slide up to 5 min at RT.

4.2.5 Time trajectories of single *Sc*.D135-L14* molecules reveal three distinct conformations in equilibrium

Characteristic time trajectories of such surface immobilized *Sc*.D135-L14* ribozymes in 10 mM and 100 mM Mg^{2+} and room temperature are shown in Figure 49. The time traces depict repeated stochastic transitions between a maximum amount of three reoccurring FRET ratios at ~0.25, ~0.4 and 0.6. Hence, the traces imply that the *Sc*.D135-L14* ribozyme adopts at least three distinguishable structural conformations in dynamic equilibrium.

The two lower FRET states are preferentially observed - especially at low Mg^{2+} concentration (Figure 49a). The highest FRET state becomes apparent only at Mg^{2+} concentrations above 15-20 mM (Figure 49b). This state with a FRET ratio of 0.6 is never populated for a long time and the few molecules that reach it return promptly to the intermediate state. Among all the observed transitions to the highest FRET state, more than 99% occur from the intermediate state (FRET ratio ~0.4). The fraction (<1%) showing transitions directly from the low FRET state is consistent with the contribution from very short resting times in the 0.4 state, which are missed because of the 33 ms time resolution of the CCD camera. This observation is thus direct evidence that the 0.4 FRET state is a possible obligatory intermediate along the folding pathway devoid of any kinetic traps.

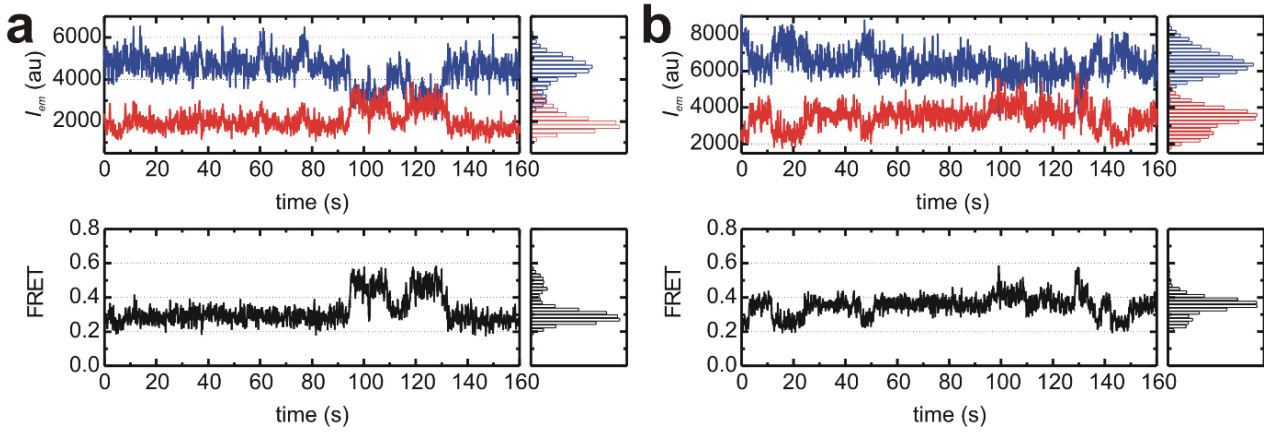


Figure 49. Single *Sc*.D135-L14* molecule time trajectories. Emission signals (top) from donor (blue) and acceptor (red) and calculated FRET efficiency (bottom) with distribution histograms (right) is shown for one molecule each at **a** 10 mM $MgCl_2$, **b** 100 mM $MgCl_2$. Traces are recorded in 80 mM MOPS, pH 7.0, 500 mM KCl at RT.

4.2.6 Magnesium ions affect single molecule folding of *Sc*.D135-L14*

Mg^{2+} ions affect the relative stability of each conformation, as evidenced by the FRET time trajectories and the dwell time spent in each state. At 10 mM Mg^{2+} and below, molecules stay predominantly in the low FRET state (Figure 49a), while at 100 mM Mg^{2+} the intermediate state becomes dominant (Figure 49b). In order to quantify the effect of Mg^{2+} on the stability of each state, FRET distributions from typically 150 single molecule trajectories at Mg^{2+} concentrations ranging from 1-100 mM were calculated (Figure 50). At 1 mM Mg^{2+} the *Sc*.D135-L14* ribozyme resides exclusively in the low FRET state and no structural dynamics are observed. This state is easily distinguished from complexes lacking the acceptor fluorophore (caused by photobleaching or blinking), because the observed FRET ratio for the latter molecules is considerably lowered to ~0.12 (Figure 50, top). There is a significant amount of molecules that remain static in the 0.25 state even at elevated Mg^{2+} concentrations (see below). Yet for analysis to calculate FRET distributions at $[Mg^{2+}] > 1$ mM, only time traces showing dynamics were considered. The distributions for the intermediate and high FRET states become apparent only above 10 and 20 mM Mg^{2+}, respectively (Figure 50). The FRET state at 0.6 is never populated for long times, which is reflected as a low amplitude distribution in the histograms. The amplitude of the intermediate FRET state distribution increases from 0.04 to 0.08 with increasing Mg^{2+} concentration, while the amplitude of the low FRET decreases from 0.23 to 0.06. This result indicates that the relative stability of the intermediate state increases concomitantly to the decrease of the low FRET state. This anti-parallel behavior, however, is not very dramatic because both states are present between 10 and 100 mM Mg^{2+}. The 0.6 FRET state distribution amplitude remains constant above ~50 mM Mg^{2+} (see also below).

Closer inspection of the histograms in Figure 50 reveals that the peak distributions of the low and intermediate FRET states shift to lower FRET values with increasing Mg^{2+}. The low FRET state distribution shifts only slightly from about 0.27 (1 mM Mg^{2+}) to 0.26 (10 mM Mg^{2+}), whereas the intermediate state distribution shifts from 0.45 (10 Mg^{2+}) to 0.40 (≥40 mM Mg^{2+}). At low Mg^{2+} concentrations (1 mM), both the states at 0.25 and 0.4 are possibly in fast dynamic equilibrium (lifetime < 30 ms) and as a result we observe only a weighted average FRET ratio a little higher than 0.25. The same could hold true for the states at 0.4 and 0.6 up to 10-15 mM $MgCl_2$, giving rise to the peak at 0.45 that shifts to 0.4 at 20 mM Alternatively, the these two FRET states at 0.25 and 0.4 each become structurally more compact above 1 mM and 20 mM Mg^{2+}, respectively, indicating decreasing distances between the two fluorophores. Neither of the two hypotheses can be ruled out at the moment, but this

result supports an idea that the group II intron conformations are connected by relatively small activation barriers.

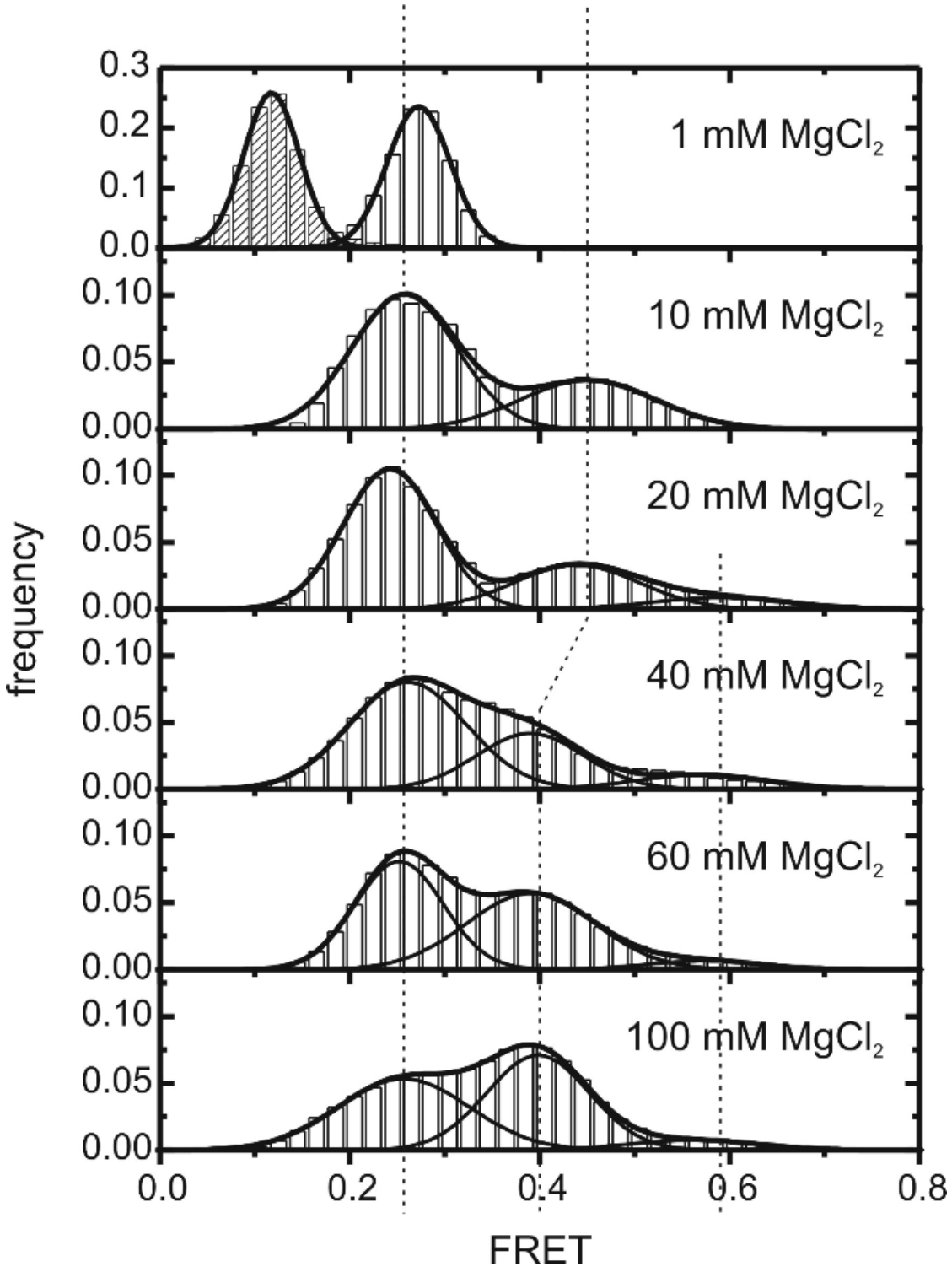


Figure 50. Histograms of FRET efficiency values from single molecule traces in dependence of Mg^{2+} concentration. All traces (constant and dynamic) were used to build the histogram at 1 mM, whereas all other histograms were created from traces showing dynamics only. Three distributions occur at FRET~0.25, ~0.4 and ~0.6. The peaks of the 0.25 and 0.4 states show a slight shift to lower FRET with higher Mg^{2+} concentration. The 0.6 distribution arises above 15-20 mM and is only minimally but steadily populated. Gaussian fits for the individual as well as the sum of all distributions are shown as solid lines. The dashed bars at 0.12 FRET corresponds to the distribution in the absence of acceptor (100 mM $MgCl_2$ background, top). Vertical lines follow the distribution maxima of the individual states at ~0.25 and ~0.6 and indicate a slight shift in the low FRET value from 1 mM compared to higher Mg^{2+} concentrations and a larger shift at the intermediate FRET state from 0.45 to 0.4.

4.2.7 The presence of Mg^{2+} increases structural dynamics

It is generally assumed that the addition of Mg^{2+} to RNA initiates folding and leads to a compact and more rigid structure that stably accumulates during the folding process. The smFRET experiments presented here reveal an opposite and so far unprecedented behavior. Although Mg^{2+} clearly initiates folding of *Sc*.D135-L14* leading to the collapsed 0.25, 0.4 and 0.6 states, the dynamics of the single molecules considerably increases with higher amounts of Mg^{2+} present. At very low Mg^{2+} the molecules remain in the 0.25 FRET state, but at higher concentrations an increasing number of transitions between the three states is observed. When plotting the fraction of dynamic molecules versus Mg^{2+} concentration, a midpoint K_{Mg} = 41.4 ± 1.4 mM is calculated (Figure 51a) corresponding again very well to the K_{Mg} for bulk-folding and cleavage (Figure 47a, Table 8). Even at high Mg^{2+} background, the fraction of dynamic molecules does not exceed 50%, being in apparent discrepancy to the cleavage experiments of the same *Sc*.D135-L14* construct (Figure 44). However, the cleavage assays are conducted at 42 °C, whereas the smFRET experiments are done at 22 °C for technical reasons. This suggests that the higher temperature is needed to enable the remaining molecules to reach the active conformation. Taken together, these two results clearly link the dynamics of *Sc*.D135-L14 to its catalysis, and raise the interesting possibility that domain motion contributes to successful reactivity.

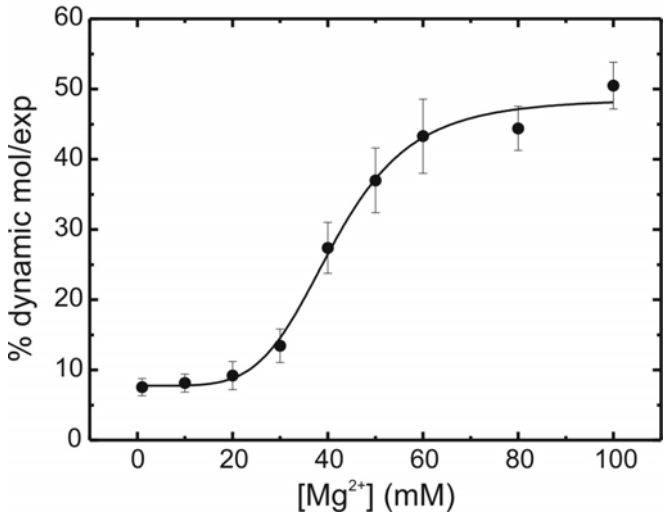


Figure 51. Mg^{2+} dependence of single molecule dynamics. Upon Mg^{2+} addition, the ratio of dynamic molecules increases from 7 to 49% with K_{Mg} = 41.4 ±1.4 mM

To further illustrate the Mg^{2+} requirement for these structural dynamics, the folding of single *Sc*.D135-L14* ribozymes was monitored while changing the buffer conditions from 10 to 100 mM Mg^{2+} at RT (Figure 52). During the first 50 s, the ribozyme is kept in 10 mM Mg^{2+} and displays a FRET ratio around 0.25 that matches the FRET histogram obtained under steady state conditions (Figure 49a, bottom right). Between 50 and 58 seconds the buffer on the microscope slide was exchanged using a syringe pump to increase the Mg^{2+} concentration to 100 mM. After a dead time of ~10 seconds, the FRET ratio jumps to 0.4 indicating that a Mg^{2+}-induced conformational change to the intermediate state takes place when the Mg^{2+} concentration is increased. The ribozyme also becomes dynamic as the FRET ratio

interchanges between the 0.25 and 0.4 FRET states. The FRET histogram corresponding to the high Mg^{2+} period also matches the FRET histogram obtained in steady state conditions (Figure 49b, bottom right). This experiment clearly shows that the individual FRET states are directly dependent on the Mg^{2+} concentration, and that the *Sc*.D135-L14* ribozyme readily reacts to an increased Mg^{2+} background by reaching a higher FRET state.

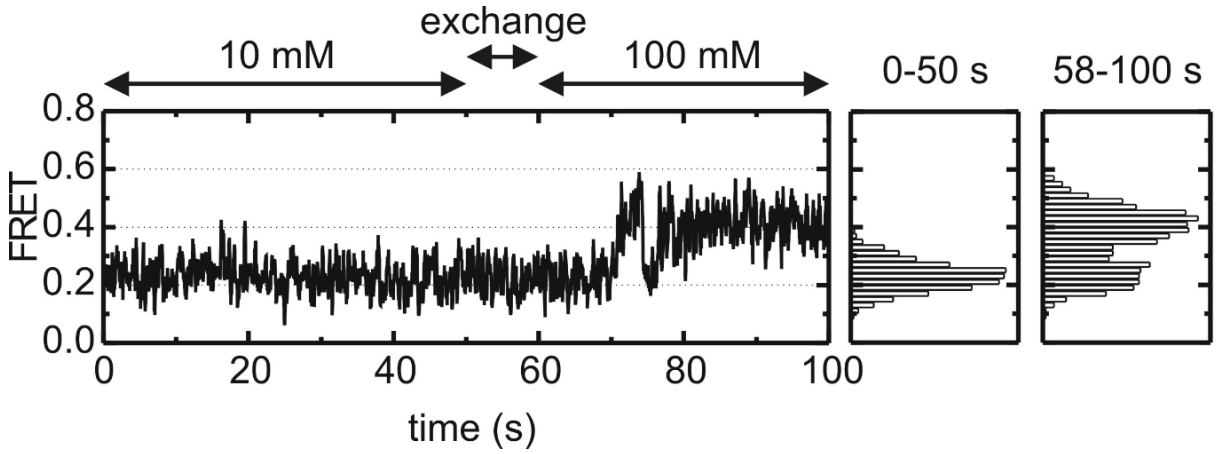


Figure 52. Dynamics occur with high Mg^{2+} concentration. The smFRET trace depicts a dynamic behavior only upon increase in $MgCl_2$ concentration from 10 to 100 at 50-58 seconds. The distributions of the states before and after addition of Mg^{2+} are shown to the right.

4.2.8 Assignment of smFRET states to conformational states reveals a so far unprecedented intermediate

The three distinguishable smFRET states at 0.25, 0.4 and 0.6 correspond to conformational states of *Sc*.D135-L14*. In a first attempt it seems obvious to assign the smFRET states to the three conformations U, I and N in the folding pathway stated in literature (Figure 36b).[84] Yet, this assignment could not be verified by the smFRET results and is reconsidered. There are strong hints toward the assumption that the 0.25 state already represents the I state:

- First, minimal FRET for U and low FRET for I is expected based on the hydrodynamic radius data[130] and the distant position of the fluorophores on the primary sequence of *Sc*.D135-L14*. This assumption is confirmed by the mean distance determination of d = 136 Å by trFRET experiments at 0 mM $MgCl_2$ (Figure 48b), where exclusively U is present. Minimal FRET in smFRET experiments is found to show a ratio of 0.12 as measured in control experiments in the absence of acceptor (Figure 50, dense histogram on top). The low FRET state in experiments with fully labeled *Sc*.D135-L14* on the other hand shows a FRET ratio of 0.25 and can clearly

be distinguished from molecules lacking acceptor or adopting a very large distance (>100 Å, Figure 16c, Chapter 1.8.3) between the fluorophores, i.e. from U.

- Second, only a few molecules are expected to be in the U state at $[Mg^{2+}] > 20$ mM, since *Sc*.D135-L14* is incubated with Mg^{2+} for 15 min at 42 °C prior to immobilization and smFRET measurements and overall Mg dependencies of folding and cleavage are found to be in the order of K_{Mg} = 20-40 mM. The molecules fold to I (and beyond) under these conditions and the U state therefore theoretically completely disappears at high Mg^{2+} concentrations, whereas in the experiments the 0.25 state is the most abundant state up to 60-80 mM Mg^{2+} (Figure 50).
- Third, the I → U transition is described to be very slow[84] and certainly is much slower than the duration of the smFRET time traces (> 200 s, Figure 49). Therefore, no such unfolding processes would be observable, whereas the smFRET traces show multiple jumps between the 0.25 and 0.4 states.

The folding pathway of *Sc*.D135-L14* is therefore refined and extended by a further on-pathway folding intermediate since there is evidence for three structural conformations at FRET rations 0.25, 0.4 and 0.6 in dynamic equilibrium starting from I (Figure 53). The folding pathway begins at the unfolded state (U) and proceeds through the extended intermediate (I) and the folded intermediate (F) to the native state (N) and reveals an intermediate state that has not been described yet.

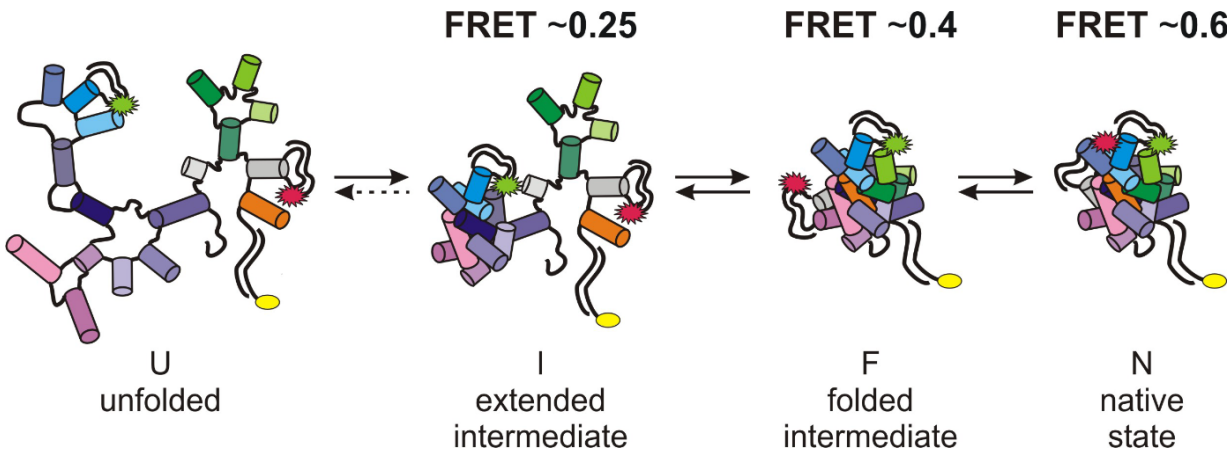


Figure 53. Refined folding pathway of *Sc*.D135-L14*. The pathway starts from the unfolded state U and in a first step, the extended intermediate I is formed upon compaction of D1. The remaining domains then pack onto D1 to give the folded intermediate F, which - only in small amounts - rearranges to the native state N in a last step. The FRET ratios of the individual states are shown on top of the cartoon. The very slow transition from I → U is depicted as a dashed arrow. Cy3 and Cy5 are depicted as green and red markers, respectively, and the biotin label is shown as a yellow ellipse. The domains are colored in shades of blue-purple for D1, light gray for D2, shades of green for D3, gray for D4 and orange for D5. DNA-oligos and major single stranded region on *Sc*.D135-L14* are depicted as black lines.

These three states can not be resolved in bulk FRET experiments because these experiments only yield ensemble-averaged results. The contribution of a short-lived and unstable conformation is weak and hidden and simply disappears in experiments that measure overall signals of a lot of molecules at the same time. Furthermore, the single molecule results reveal that a large fraction of the molecules remain static in the low FRET state even at high Mg^{2+} concentrations and explain the observed overall low FRET ratio in bulk (Figure 47a).

4.2.9 Dwell time analysis determines the rates of folding

In order to quantify the relative stability of the three conformations observed in the single molecule experiments, the rates for folding of *Sc*.D135-L14* were determined. The waiting times (τ) in each state between two conformational transitions from about 150 single molecule time trajectories were measured. These lifetimes between the transitions from I to F and back were termed τ_1 and τ_{-1}, respectively (Figure 54a). Analogously, the lifetimes τ_2 and τ_{-2} describe the resting times in the F state before going to N and vice versa, respectively, dividing the dwell times in the F state into two categories τ_{-1} and τ_{-2} depending on the final state (I or N, respectively).

The lifetime distribution of each state was calculated and a characteristic example is shown for τ_1 at 50 mM Mg^{2+} in Figure 54b. The distribution decays exponentially as expected for a stochastic process. To obtain folding rates k_1 and k_{-1} from the respective lifetimes τ_1 and τ_{-1}, the distributions were integrated into a cumulative time distribution (Figure 54c), which is then fit to an exponential growth curve as described.[140, 213, 261] The integrated distribution reveals folding heterogeneity in the form of two distinguishable exponentials. Folding heterogeneity has been observed in numerous single molecule studies on other ribozymes.[213, 234, 262, 263] The major component for k_1 exhibits an exponential rate k_1 = 0.66 s^{-1} at 100 mM $MgCl_2$. The folding rate k_{-1} = 0.51 s^{-1} was obtained similarly, also observing folding heterogeneity. The measured rates are in good agreement with the FRET histogram distributions in Figure 50 (bottom). At 100 mM Mg^{2+}, the ratio k_1/k_{-1} = 1.3 corresponds well to the ratio of distribution amplitudes for the F and I states, $A_{0.4}/A_{0.2}$ = 0.78/0.56 = 1.4.

The number of transitions observed from the F to the N state is not statistically significant to determine k_2 and k_{-2} unambiguously, because k_2 is slower than k_{-1} and the highest FRET conformation is only rarely reached. To estimate the value of these folding rate constants, the average dwell time in the F state before jumping to the N state and in the N state before

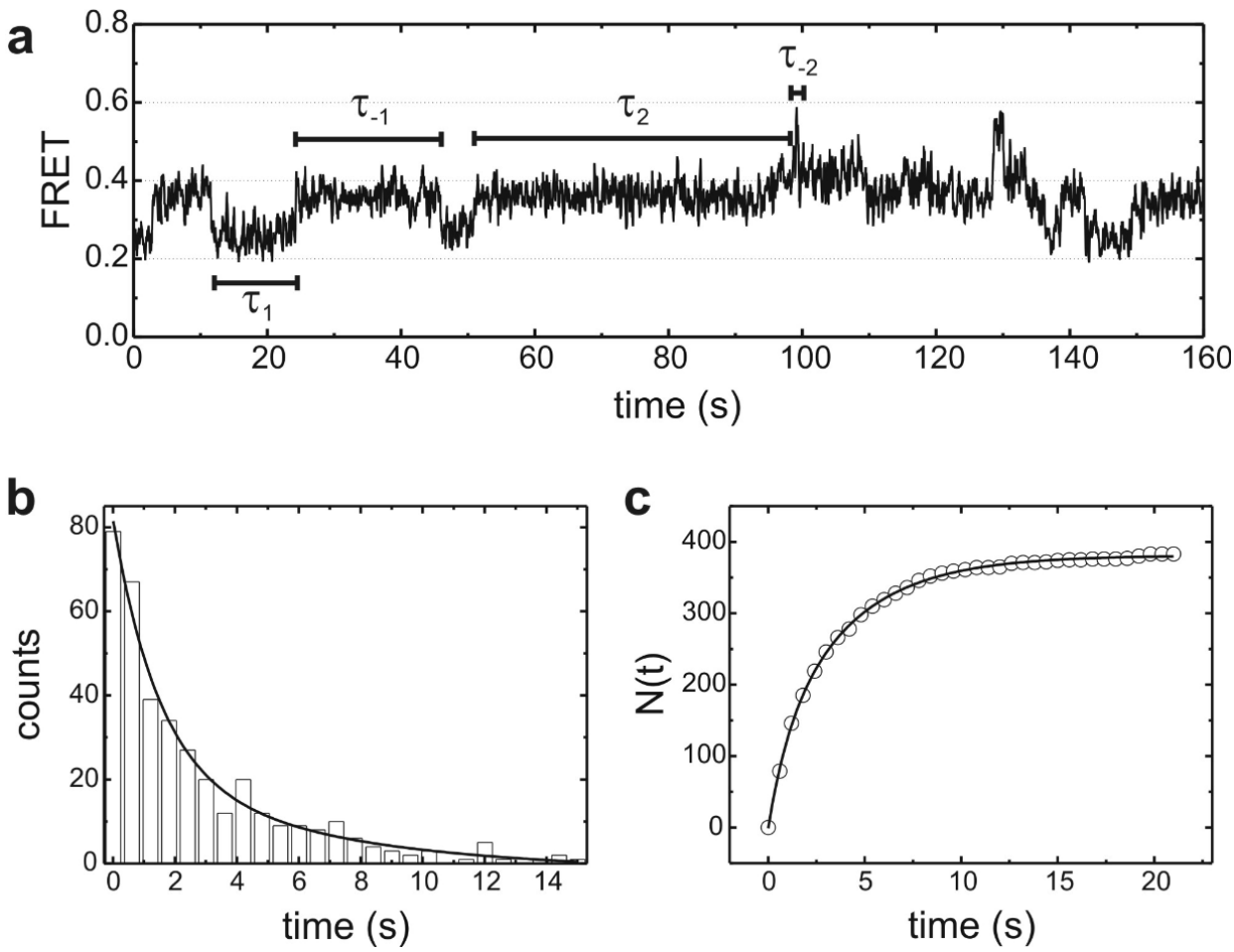


Figure 54. Dwell time analysis for calculation of folding rates. **a** Single molecule trajectory is divided into dwell times that are terminated by transition to neighboring conformational states. The classification of dwell times is indicated and shows the assignment of τ_1- τ_4 dependent on the starting and the final conformational state. **b** Representative histogram of dwell times in the 0.25 FRET state (τ_1) at 50 mM Mg^{2+} showing the number of events in a particular dwell time range. The solid line corresponds to the double exponential decay fit characterizing the rate k_1. **c** The number of dwell times N(t) in the low FRET state shorter than time t obtained by integration of the dwell time histogram shown in b. The double exponential growth fit is shown as solid line.

coming back to F were calculated for ~18 molecules (τ_{av} = 15.3 ± 14.7 s and 0.6 ± 0.3 s, respectively) and result in k_2 = 0.07 ± 1.0 s^{-1} and k_{-2} = 1.7 ± 0.7 s^{-1} as the inverse of these values. At 100 mM Mg^{2+}, the ratio of the slow rate k_2 and the fast k_{-2} (k_2/k_{-2} = 0.04) is in reasonable agreement with the ratio of the relative amplitudes of the 0.6 and 0.4 FRET distributions ($A_{0.6}/A_{0.4}$ = 0.008/0.078 = 0.10, Figure 50) and confirms the low abundance of the N state in the FRET histogram.

As mentioned above, a significant fraction of the molecules remains static in the I state (Figure 51). Since these traces do not show any confined lifetimes, they were not included in the dwell time analysis.

4.2.10 The forward rates are independent of Mg^{2+}

The folding rates k_1 and k_{-1} were determined as described above in the Mg^{2+} concentration range of 5-100 mM (Figure 55). Below 5 mM Mg^{2+}, the *Sc*.D135-L14* ribozymes become static in the I state and thus not enough dynamic events are observed to quantify the rates with confidence. Evidence for folding heterogeneity is detected for k_1 and k_{-1} at all concentrations (Figure 56, Table 10). At 5 mM Mg^{2+}, the major component of k_{-1} is 1.5 fold larger than the major component of k_1, which is consistent with I being more stable than F at low Mg^{2+} (Figure 50). k_1 remains approximately constant, whereas k_{-1} first rises and then drops at Mg^{2+} concentrations higher than 40 mM, shifting the equilibrium from the I towards the F state (Table 10). This is again consistent with the change in relative stability between I and F as observed in the FRET histogram distributions (Figure 50). The rate constants k_2 and k_{-2} are estimated from average dwell times of a few molecules showing that both rates remain approximately constant around 0.07 and 1.3 s^{-1}, respectively.

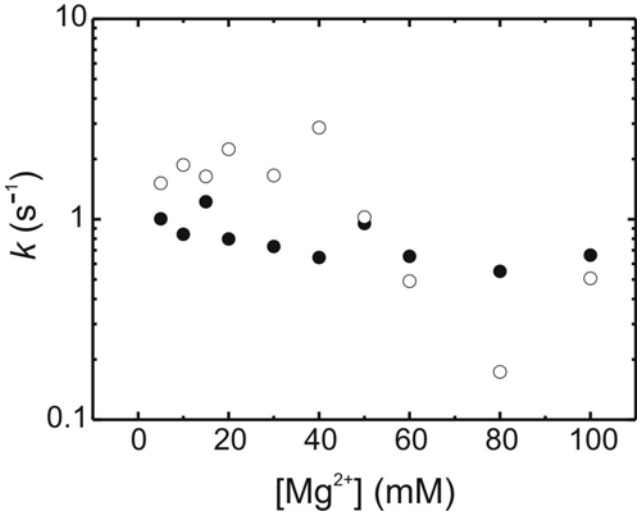


Figure 55. Effect of Mg^{2+} concentration on the observed single molecule folding rates k_1 (●) and k_{-1} (○).

Folding rates independent of Mg^{2+} concentration have been described for the C-domain of RNase P[234] and the *Tetrahymena* ribozyme.[262] For *Sc*.D135, Waldsich and Pyle have recently shown that the rate of D1 compaction is independent, whereas the amplitude of compaction is dependent of Mg^{2+} concentration.[225] Such a behavior is explained by a conformational capture mechanism, in which the correct conformation (of the κ-ζ region) is unstable and must be trapped by Mg^{2+} coordination. Based on the results presented in this work, this Mg^{2+}-independence of *Sc*.D135 compaction progresses along the folding pathway of this ribozyme. The distribution analysis of the different states reveals further that the relative values of the rates are not only consistent with the FRET histogram distributions, but also that also the amplitudes are directly proportional to the Mg^{2+} concentration (Figure 50). The rise in amplitude of the collapsed *Sc*.D135 (F) with increasing Mg^{2+} concentration is consistent with results from hydroxyl radical footprinting, where D5 becomes fully protected only above 100 mM Mg^{2+}.[125] Taken together, this indicates that to capture D3 and D5, as well as to stabilize the *Sc*.D135 complex, Mg^{2+} ions are the determining factor under *in vitro*

conditions. Mg^{2+} ions could thereby either serve to overcome electrostatic repulsion between the negatively charged domains or to stabilize specific tertiary contacts.

Table 10. Rates k_1 and k_{-1} and amplitudes A_1 and A_{-1} from double exponential fit at tested Mg^{2+} concentration. Values are corrected as described in Supplementary Text to [254]. The error limits correspond to one standard deviation. Folding rates are given in s^{-1}.

$[Mg^{2+}]$ (mM)	$k_1(1)$	$A_1(1)$	$k_1(2)$	$A_1(2)$
5	1.00 ± 0.04	125 ± 7	0.13 ± 0.01	94 ± 1
10	0.84 ± 0.06	148 ± 13	0.13 ± 0.02	76 ± 2
15	1.22 ± 0.04	170 ± 9	0.11 ± 0.01	106 ± 1
20	0.80 ± 0.06	111 ± 11	0.11 ± 0.02	88 ± 2
30	0.73 ± 0.07	64 ± 7	0.09 ± 0.02	48 ± 1
40	0.65 ± 0.06	95 ± 11	0.14 ± 0.03	49 ± 3
50	0.95 ± 0.15	115 ± 31	0.24 ± 0.01	285 ± 8
60	0.65 ± 0.05	120 ± 11	0.11 ± 0.02	72 ± 2
80	0.55 ± 0.05	108 ± 8	0.18 ± 0.00	65 ± 1
100	0.66 ± 0.06	434 ± 34	0.06 ± 0.04	256 ± 11

$[Mg^{2+}]$ (mM)	$k_{-1}(1)$	$A_{-1}(1)$	$k_{-1}(2)$	$A_{-1}(2)$
5	1.52 ± 0.09	131 ± 19	0.13 ± 0.03	100 ± 3
10	1.87 ± 0.09	128 ± 16	0.16 ± 0.01	120 ± 1
15	1.64 ± 0.02	274 ± 10	0.10 ± 0.01	43 ± 1
20	2.24 ± 0.08	166 ± 18	0.12 ± 0.01	82 ± 1
30	1.66 ± 0.05	85 ± 6	0.11 ± 0.01	46 ± 0
40	2.87 ± 0.14	60 ± 12	0.08 ± 0.00	120 ± 1
50	1.02 ± 0.08	269 ± 38	0.14 ± 0.03	154 ± 6
60	0.49 ± 0.04	69 ± 5	0.07 ± 0.01	151 ± 1
80	0.17 ± 0.03	145 ± 12	0.08 ± 0.00	58 ± 4
100	0.51 ± 0.05	148 ± 14	0.09 ± 0.01	272 ± 4

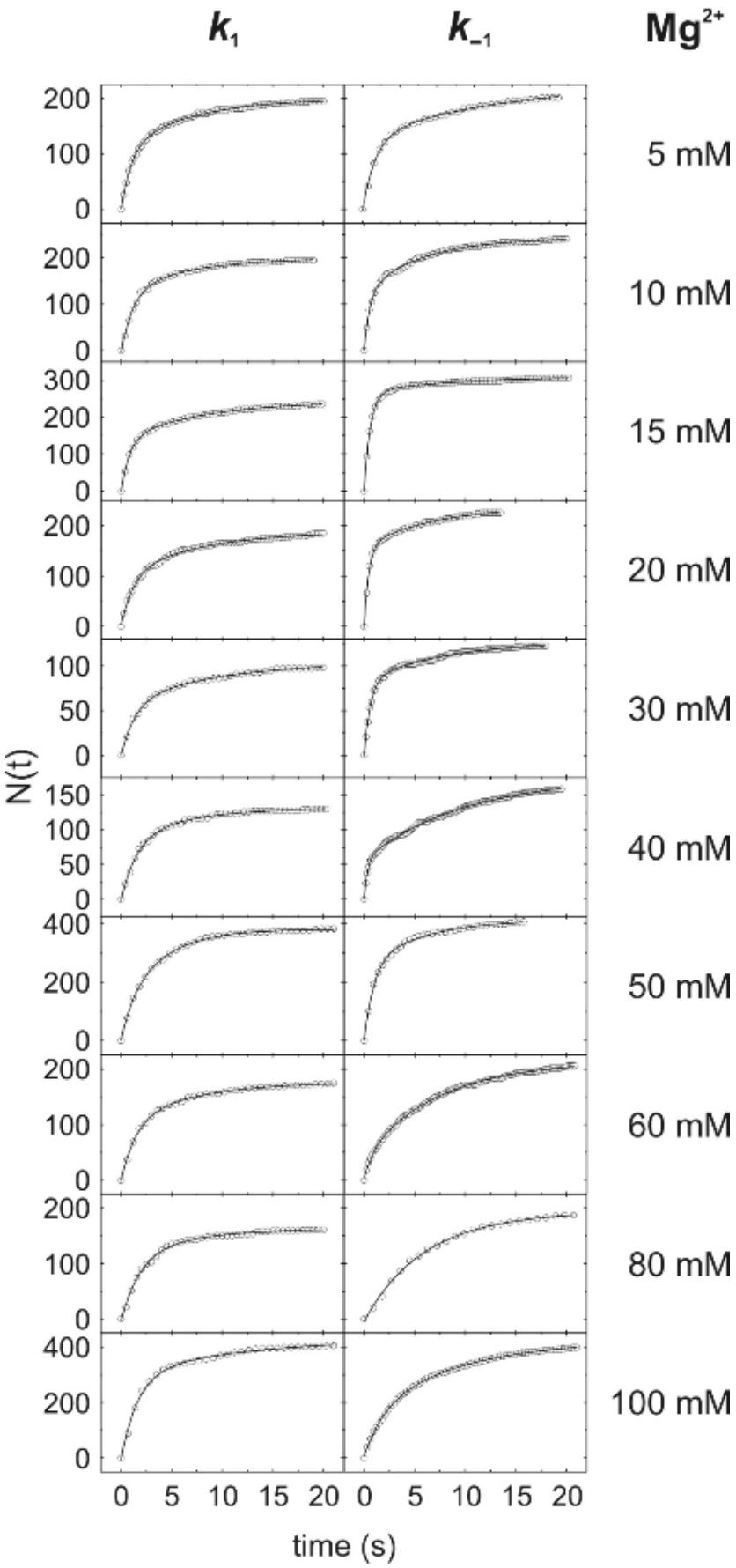


Figure 56. Calculation of rates k_1 (left column) and k_{-1} (right column). Number of dwell times shorter than time t obtained form integration of dwell time histograms at $MgCl_2$ concentrations as indicated to the right are shown together with double exponential fits (black lines).

4.2.11 Substrate binding stabilizes the 0.6 FRET conformation and leads to the according changes in folding rates

In our FRET studies, the F and the N states become apparent as two distinct distributions above 20 mM Mg^{2+}. The presence of these two more compact distributions thereby coincides with the onset of catalytic activity (compare Figures 45, 47 and Figure 50), suggesting that one of these two conformations corresponds to the catalytically active structure. However, it has been suggested that *Sc*.D135 collapses directly to the active state with a midpoint of about 20-40 mM Mg^{2+}.[125] The question now is: Which of the states with higher FRET represents the active conformation?

In order to elucidate the active conformation, smFRET studies on *Sc*.D135-L14* in the presence of the substrates 17/7 and 17/7-dC were performed. The retarded cleavage of 17/7-dC (k_{obs} = 0.05 ± 0.01 min^{-1}, Table 7) enables the observation of the substrate bound

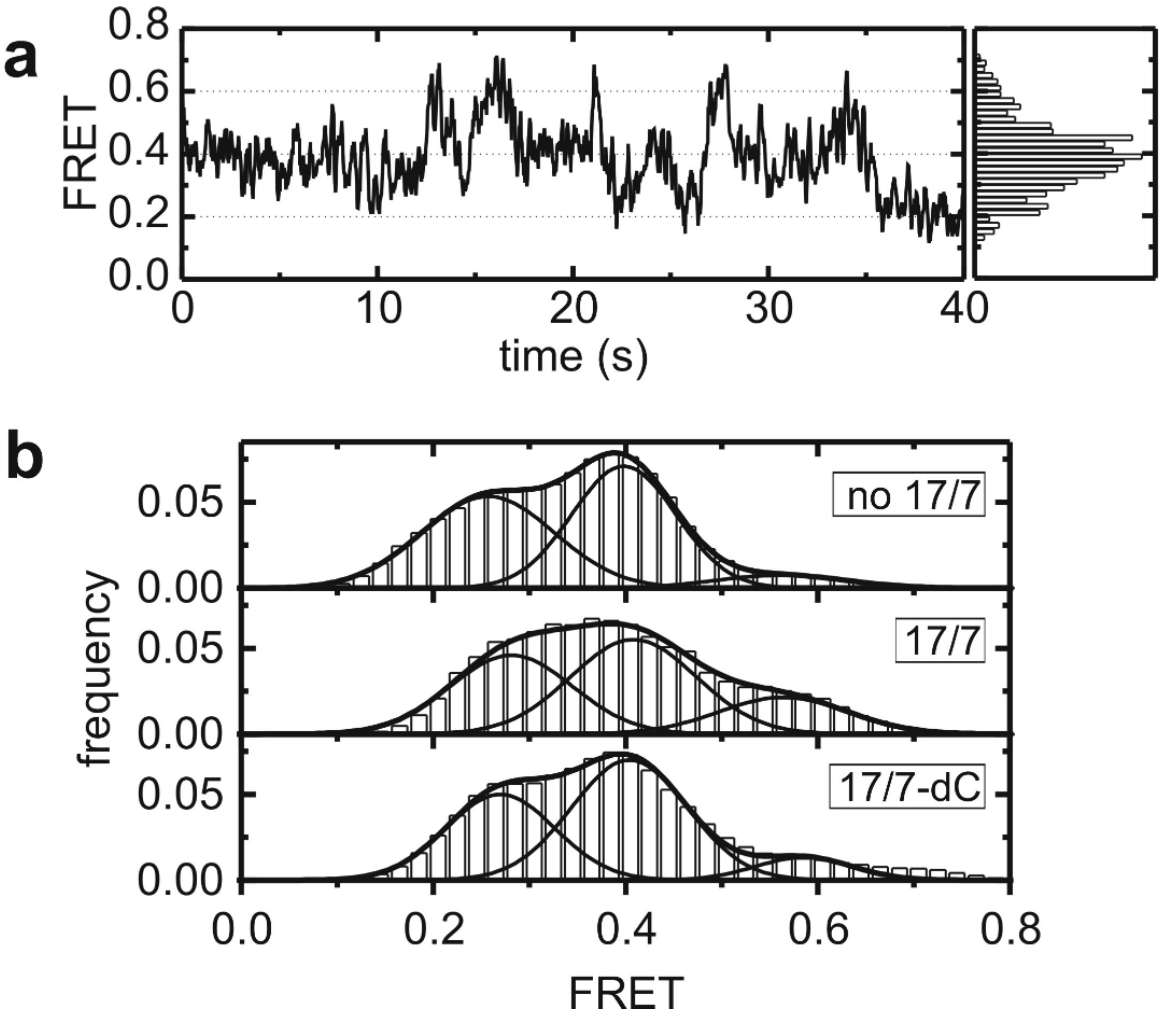


Figure 57. The presence of substrate effects *Sc*.D135-L14* dynamics and indicates native state stabilization. **a** Characteristic single molecule trajectory observable in the presence of 17/7-dC. The 0.6 state occurs at high frequency (compare to Figure 14B and 15). Distribution is shown to the right. **b** Histogram of FRET efficiency in the absence of substrate (top), in the presence of substrate 17/7 (middle) and in the presence of the slower cleaving substrate 17/7-dC (bottom). The 0.6 state is clearly stabilized by substrate binding.

Sc.D135-L14* molecules for longer periods of time. The presence of either 17/7 or 17/7-dC clearly stabilizes the N state as indicated by the stronger contribution of the 0.6 FRET to the histograms (Figure 57), suggesting that the high FRET state corresponds to the active conformation. Consequently, the intermediate state represents a hitherto undetected obligatory intermediate, or near-native state, from which only few molecules reach the active state at a time. Interestingly, the substrate 17/7 leads to a higher occurrence of the 0.6 state compared to the 17/7-dC oligonucleotide. A plausible explanation for this finding could be that the 2'-OH at the cleavage site makes contacts that contribute to the assembly of the catalytic core and hence, also to the overall architecture.

Additionally, the folding rates k_2 and k_{-2} were measured in the presence of 17/7 and 17/7-dC from 6 and 17 single molecule trajectories, respectively. Substrate binding increases k_2 by one order of magnitude (1.3 ± 1.8 and 0.7 ± 1.2 s^{-1}, respectively) and k_{-2} decreases by 30-50% (0.8 ± 1.1 and 1.2 ± 1.8 s^{-1}, respectively), further indicating that the substrate stabilizes folding into the N state, in agreement with the above described findings.

4.2.12 Ca^{2+} changes the folding rates and results in accumulation of *Sc*.D135-L14* in an altered 0.6 state

In a *trans* cleavage assay using exD123 and D5, it was found that Ca^{2+} severely inhibits the first step of selfsplicing (Figure 58).[25, 201] No cleavage product was obtained in 100 mM Ca^{2+} or a mixture of 100 mM Mg^{2+} : 100 mM Ca^{2+}. A series of assays at lower Ca^{2+} concentrations, i.e. 1-5 mM $CaCl_2$, was then carried out. In these assays, $MgCl_2$ was added to obtain a total divalent cation concentration of 100 mM in order to work under conditions of the same ionic strength and reliable comparison of the kinetic profile of reactions at different Ca^{2+} concentrations. Under all conditions, the reactions still obeyed pseudo-first order kinetics.

The cleavage rates were found to decrease roughly 2-fold upon addition of 5 mM Ca^{2+} (k_{obs} = 0.023 ± 0.01 min^{-1}) compared to the reaction in Mg^{2+} only (k_{obs} = 0.045 ± 0.02 min^{-1}) (Figure 58).[25] Further addition of Ca^{2+} continuously decreases the reaction rate and at a ratio of 80:20 for Mg^{2+}:Ca^{2+}, no splicing product was detected after 20 min incubation at 42 °C. Additionally to these reactions that were performed in mixed metal ion solutions from the beginning of the reaction, Ca^{2+} was also found to inhibit the formation of the products if Ca^{2+} was added to a reaction conducted exclusively in Mg^{2+} at a later point in time during the reaction.[25] Taken together, these results indicate that Ca^{2+} either disrupts the active three-dimensional structure or directly interferes with catalysis. Additionally, Ca^{2+} is able to

actively replace Mg^{2+} in a prefolded state of the molecules from at least one of its crucial binding pockets and consequently, to inhibit splicing.[25]

Knowing about this inhibitory effect of Ca^{2+} in cleavage reactions, smFRET experiments on *Sc*.D135-L14* were performed to further characterize the impact of Ca^{2+} on folding by analyzing single molecule time traces. Experiments were conducted in the absence of substrates to determine the influence of Ca^{2+} on folding independently of catalysis. Analogous to the cleavage reactions described above, the amount of Ca^{2+} was increased from 0 to 2, 5, 7 or up to 10 mM in a background of Mg^{2+} to a total divalent cation concentration of 100 mM. Distribution histograms of the resulting FRET ratios in single molecule time trajectories are shown in Figure 59. The contribution of the 0.6 state is dramatically enhanced at Ca^{2+} concentration of 7 and 10 mM! On the other hand, the magnitude of the 0.4 state decreases as Ca^{2+} is increased. Careful inspection of the histograms reveals a shift in maximum at the 0.6 state from 0.57 (0, 2, 5 mM Ca^{2+}) to 0.55 (7 mM Ca^{2+}) and 0.54 (10 mM Ca^{2+}), whereas the peaks at 0.25 and 0.4 stay constant.

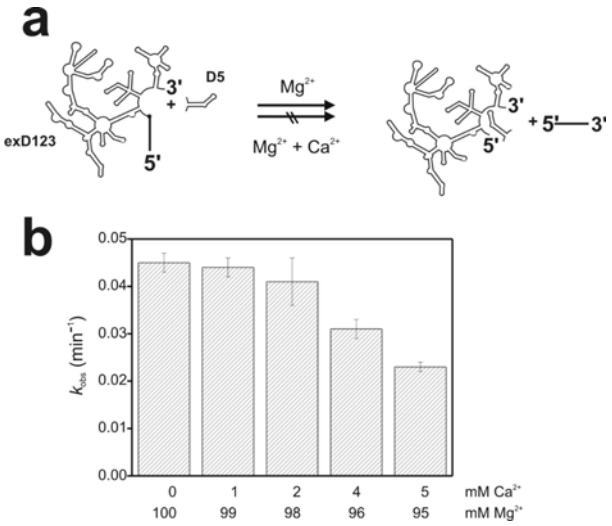


Figure 58. *Trans* cleavage reaction of exD123 (substrate) and D5 (catalyst) is inhibited by increasing concentrations of Ca^{2+}. **a** The cleavage reaction of exD123 and D5, resulting in the release of the exon (ex) piece, is shown schematically. **b** A concentration of only 5 mM Ca^{2+} in a 95 mM Mg^{2+} background already results in a roughly two-fold slowdown of the cleavage rate. Figures are adapted from [25].

Dwell time analysis was performed as described for experiments with changing Mg^{2+} concentrations (1-100 mM Mg^{2+}, Section 3.2.9) and at 100 mM Mg^{2+} in the presence of substrates (Section 3.2.11): Dwell time distributions of typically 30-50 molecules were integrated and fitted with exponentials, revealing heterogeneity and determining the folding rates k_1 and k_{-1}. The folding rates k_2 and k_{-2} were calculated from averaged dwell times. The obtained folding rates (Table 11) reveal a dependency on Ca^{2+} concentration: k_1 and k_2

Table 11. Folding rates in the presence of various concentrations of Ca^{2+}.

Ca^{2+} [mM]	Mg^{2+} [mM]	k_1 (s^{-1})	k_{-1} (s^{-1})	k_2 (s^{-1})	k_{-2} (s^{-1})	# mol
0	100	0.66 ± 0.06	0.51 ± 0.05	0.07	1.70	253/18
2	98	0.78 ± 0.22	0.13 ± 0.02	0.18	1.08	48
5	95	0.65 ± 0.13	0.06 ± 0.01	0.26	0.60	33
7	93	1.53 ± 0.18	0.25 ± 0.01	0.47	0.63	35
10	90	1.53 ± 0.26	0.20 ± 0.02	0.78	0.29	50

increase and k_{-1}, k_{-2} decrease, each one showing an approximately linear, but individually distinct tendency (Figure 60).

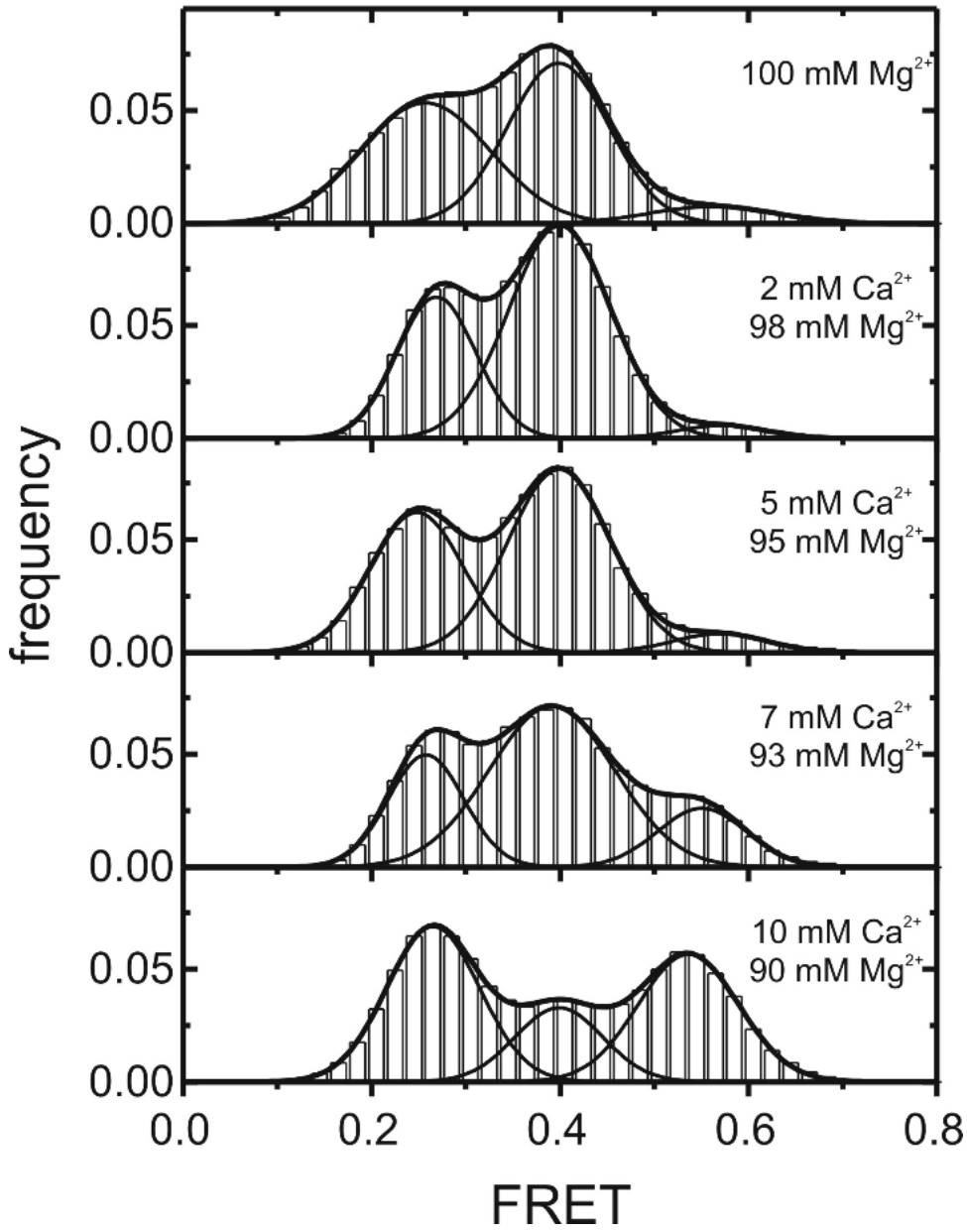


Figure 59. Distribution histograms of smFRET time traces in the presence of Ca^{2+}. Upon addition of up to 10 mM Ca^{2+} (in a total divalent metal ion concentration of 100 mM completed with the addition of Mg^{2+}), the contribution of the high FRET state at ~0.6 dramatically increases. Concomitantly, the mean maximal FRET value shifts from 0.57 to 0.54 and the peak height of the 0.4 FRET state decreases.

The histograms show an accumulation of the molecules in the 0.6 FRET state. This is due to increased lifetimes in the folded 0.6 FRET state. The molecules still sparsely reach the 0.6 FRET state, but once there, they will populate it for a longer time before returning back to the 0.4 FRET state. The forward rates increase, whereas the backwards rate decrease, indicating stabilization of the transition states. The molecule apparently requires less time for folding to the next state and spends more time in these states before unfolding. Yet, this change in

folding profile results in inhibition of cleavage. Interestingly, there is a slight shift in distribution maximum of the highest FRET state at ~0.6 (N). The lower FRET states (I and F) do not experience any shift in the FRET ratio. This is a possible indication, that the high FRET state that forms in the presence of Ca^{2+} is slightly different than in the presence of Mg^{2+}. This high FRET state formed in the presence of Ca^{2+} therefore would represent an aberrant, misfolded species. In this case however, the calculated rates would not exclusively reflect the rates of formation of this misfolded species (M), but rather a mixture from dynamics to M as well as the correctly formed (active) state N since both Mg^{2+} and Ca^{2+} are present at all instances.

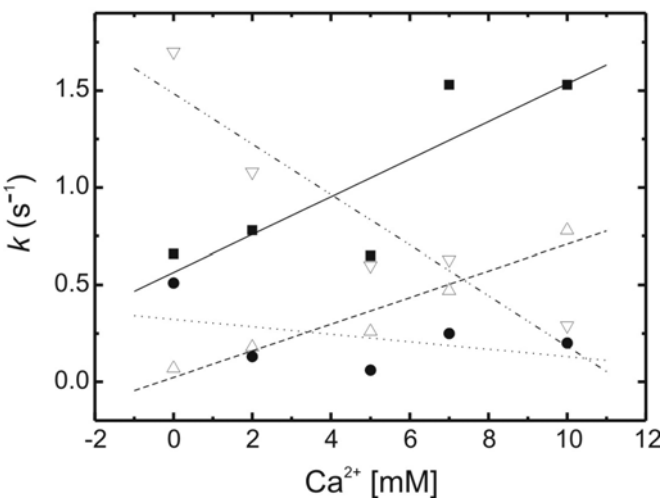


Figure 60. Folding rates in the presence of Ca^{2+}. Linear changing tendencies are reported for k_1 (■, solid line), k_{-1} (●, dotted line), k_2 (△, dashed line) and k_{-2} (▽, dash-dotted line).

However, the folding of *Sc*.D135-L14* from I to a higher FRET state is impaired in the presence of Ca^{2+} compared to Mg^{2+}. Waldsich et al. state a folding control mechanism for tertiary structure formation of *Sc*.D135,[56] in which the substructure within D1 including the κ-ζ element controls the specific collapse of the molecule. Furthermore, this distinct substructure is described to be a metal ion binding pocket.[225] It is therefore possible that the observed effect on folding is due to binding of a Ca^{2+} to the κ-ζ element and thereby, interferes with folding control. The κ-ζ element serves also as the docking platform for D5 binding later in the folding pathway.[56] A Ca^{2+} metal ion at this central core structure might therefore influence the correct formation of the active site structure and as a result, interfere with catalysis.

The utilization of the smFRET insights for interpretation of the results from the cleavage assays is not straightforward, since the RNA systems studied in the different assays are not equal. Nevertheless, the interesting finding that Ca^{2+} replaces Mg^{2+} in prefolded molecules might be connected to the fact that the molecules are highly dynamic even at high divalent metal ion concentrations. The metal ion replacement might not (only) happen in the folded state, but the binding site(s) might become accessible in the intermediate forms that are potentially populated long enough for the exchange to take place. In the subsequent folding transition(s), the high FRET state reached might be slightly altered and no longer active.

4.3 Discussion and Conclusion

A new paradigm for folding of large multidomain RNAs

The *Sc*.D135-L14* RNA presented in this work is the largest protein free ribozyme system studied so far by single molecule FRET. The folding pathway of this remarkable intron is characterized in detail and reveals a folding procedure including four steps: The folding pathway starts with the formation of U from a random coil structure, and proceeds through three subsequent folding steps with a previously uncharacterized on-pathway folding intermediate, extending earlier studies that suggested a folding pathway devoid of kinetic traps. The identification of a hitherto unknown on-pathway folding intermediate has direct implications not only for the folding pathway but also catalysis of these large molecular machines. Group II intron folding is unique but might very well represent a paradigm for the multidomain assembly of large RNAs and RNA-protein complexes. Our single molecule experiments reveal important new concepts for RNA folding in general and also explain long-standing questions for group II intron catalysis.

The slow and overall rate-limiting compaction of D1 at low Mg^{2+} concentration yields a transient intermediate I, which corresponds to the 0.25 FRET state in the experiments described in this work. The slight decrease in FRET from 0.28 to 0.24 between 1 and 20 mM could thereby reflect the capturing of Mg^{2+} by the κ-ζ region, as described in NAIM studies,[225] and the subsequent structural compaction.

With increasing Mg^{2+}, two additional states F and N at 0.4 and 0.6 FRET arise stepwise. Mg^{2+} concentration > 10 mM are needed to substantially populate F. Nevertheless, F becomes dominant only at Mg^{2+} concentrations approaching 100 mM, consistent with results from hydroxyl radical footprinting, where D5 becomes fully protected only above 100 mM Mg^{2+}.[125] The accumulation of F is paralleled by the decrease of the low FRET state I (Figure 50). In accordance with previous results,[56, 84, 225] the F state corresponds to a compact *Sc*.D135 state. Similar to the folding of the C-domain of RNase P[234] and the *Tetrahymena* ribozyme,[264] the rates for the transition between I and F are independent of Mg^{2+} (Figure 55), suggesting a conformational capture mechanism to anchor the remaining domains to the D1 scaffold. This finding is highly intriguing as already the preceding D1 compaction follows this mechanism, where the rate of D1 compaction is independent but the amplitude of compaction is dependent on Mg^{2+} concentration. It thus appears that the whole pathway of *Sc*.D135 folding *in vitro* is dominated by the trapping of inherently unstable conformations by Mg^{2+}. Taken together, this indicates that to capture D3 and D5, as well as to

stabilize the *Sc*.D135 complex, Mg^{2+} ions are the determining factor under *in vitro* conditions. The very interesting and as yet undescribed finding that *Sc*.D135 molecules show distinctively more dynamics upon the presence of higher Mg^{2+} concentrations and the existence of low energy barriers between the compact species I, F and N adds to the idea of a new paradigm of folding. Rather than to stably reach the native state, the molecule uses dynamics for accurate and flexible folding as well as attendance to the environment at all times.

The newly discovered N state at 0.6 FRET very likely corresponds to the active state. It can be distinguished from F only above 20 mM Mg^{2+}, but always remains a minor species. This explains why it has not been previously observed in bulk experiments and at the same time illustrates the power of smFRET to identify short-lived conformations otherwise hidden in ensemble-averaged experiments. The relative stability between the F and N states remains constant above 40 mM Mg^{2+}, ruling out a capture mechanism involving Mg^{2+} for the F → N transition. Instead, it is binding of the substrate that leads to a stabilization of the N state, making N the catalytically active species. The change in FRET intensity by about 0.2 must be accompanied by a major conformational change of the compact *Sc*.D135-L14*: As the IBS-EBS and ε-ε' interactions are located within D1 and their specific formation will not be reported on in the FRET assay described here, it must be the tripartite λ-λ'-λ' interaction, linking the intronic 5'-end with central nucleotides in D1 and the catalytically crucial D5,[185] that causes this major rearrangement of domains. Such a stabilization of the native state by tertiary contacts resembles the situation in the extended hammerhead ribozyme,[265, 266] although in that case the responsible interaction is distant from the cleavage site.

The presence of divalent metal ion cofactors other than Mg^{2+} induces alterations in folding and changes the conformational state N, which results in inhibition of catalysis as shown for the case of Ca^{2+}. There is a high sensitivity of *Sc*.D135 towards the identity of metal ion cofactors. This finding further adds to the idea of the existence of a highly optimized folding control element that is directed by metal ion binding and whose accuracy additionally is dependent on the correct cofactor (Mg^{2+} in the case for *Sc*.D135).

One can now speculate about additional implication of our results for the splicing mechanism of group II introns when a water molecule acts as the nucleophile. The intron folds directly to the on-pathway compact core involving domains 1, 3, and 5. In the presence of the covalently bound 5'-exon, rearrangement to the active species takes place in equilibrium. It remains unknown to what extent the active species is formed under these conditions. However, after the first step of splicing, the intron falls back to the starting

collapsed state where it remains majorily. Based from this collapsed state the intron can now reach the active conformation for the second step of splicing, i.e. the cleavage of the 3'-exon (Figure 61). Also in the presence of D6, it is potentially highly advantageous for the intron to follow a reaction mechanism, where from a central state (F in our case) the different active species can be reached, falling back to the central state after the specific cleavage reaction. In addition, the molecules are sensitive to contaminations and changing conditions and might form misfolded species upon binding of an incorrect metal ion cofactor.

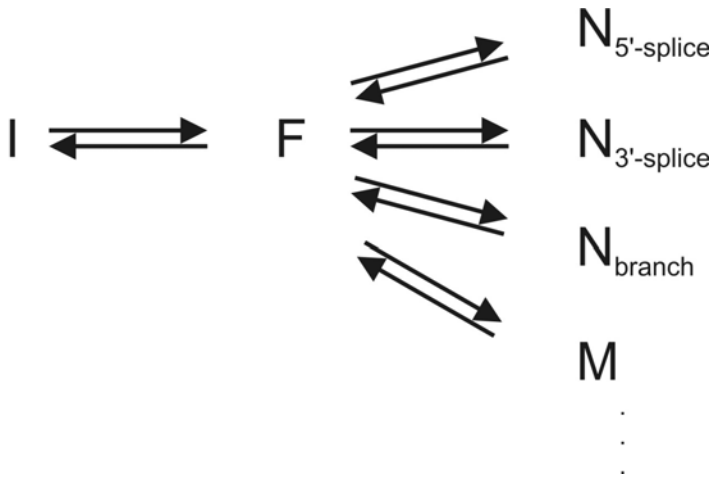


Figure 61. The folding pathway to the collapsed *Sc*.D135 state (F) as a starting point to reach the different native states for 5'-splicing (this work, $N_{5'\text{-splice}}$) and hypothetically similar native state for further catalytic steps ($N_{3'\text{-splice}}$, N_{branch}) or a misfolded, inactive species (M).

Implications for *in vivo* functioning of group II introns

Folding of a large multidomain RNA without kinetic traps following exclusively on-pathways intermediates is hitherto unknown. Nevertheless, such a highly ordered pathway is very advantageous for a retro-element like a group II intron. It allows this molecule to fold co-transcriptionally starting with D1, which is transcribed first. The slow transition from U to I thereby ensures a proper folding of this domain and subsequent docking of the downstream domains as soon as they have been transcribed (e.g. D3 and D5). After self-splicing, the ribozyme resides in the near-native state F. Indeed, the observed active state N is stabilized by substrate binding, and thus is a mimic for the first step of a retro-homing group II intron.

While *Sc*.ai5γ requires high salt and temperature *in vitro* for functioning, this ribozyme is assisted by the protein Mss116 *in vivo*.[267] Mss116 is a DEAD-box protein with ATPase as well as helicase activity.[268] It has been a matter of debate if the unwinding activity of Mss116 is needed to promote splicing [269, 270]. The fact that the folding pathway of the *Sc*.D135-L14 ribozyme is devoid of any kinetic trap on the single molecule level now supports the idea that Mss116 is only needed to stabilize the native fold at low Mg^{2+} concentration and not to help the RNA to escape from misfolded intermediates.

Due to our labeling scheme with the two fluorophores sitting in D1 and D4 and the

experimental constraints (30 ms time resolution), we are likely to miss further intermediate states, e.g. the docking of D3, as such states are formed either very fast or might not lead to large rearrangements of the RNA overall architecture and hence a considerable change in FRET using the current fluorophore positions. However, the exponential shape of the dwell time distributions does not support this idea. To conclude, group II intron folding seems to follow a unique, highly ordered pathway without any side-tracks. Over all, we still have relatively little knowledge on its energetics and driving forces, but considering the rich evolutionary heritage of these introns, it is likely that their folding pathway constitutes a new paradigm for many more large multidomain RNAs.

4.4 Additional smFRET constructs

FRET reports on the distance and the relative distance changes of two fluorophores that are fixed at certain positions within a molecule. Hence, only conformational changes that alter the positions of these two selected points relative to each other will be monitored. The domain motions detected by FRET lead to only qualitatively interpretation with respect to the two labeled domains, whereas the behavior of the other domains is only indirectly reflected in the optimal case and often remains "dark", i.e. hidden. To shine light on other aspects of the folding pathway and on conformational changes of other domains in isolation or the context of the entire molecules as well as to verify the present results, it is necessary to alter the positions of the fluorophores on the molecule.

4.4.1 GMPS transcription and labeling with Cy5-maleimide

As described in 4.1.2, it is potentially feasible to place a fluorescent dye on either of the RNA termini. A $tRNA^{Asp}$ containing a 5'-monothiophosphate was labeled postranscriptionally with fluorescein and was successfully used in FRET experiments.[244] In a similar approach, *in vitro* transcriptions of *Sc*.D135-L14 and *Sc*.D1-L1 (*vide infra*) in the presence of GMPS and subsequent attempts to label the RNAs with Cy5-maleimide were carried out to eventually result in the desired smFRET constructs (Figure 62). The *Sc*.D1-L1 RNA consists of D1 including the loop in d2b to be labeled by Cy3-DNA and the construct was obtained by introduction of a *Eco*RI restriction site after the D1 sequence into the plasmid pT7D135-L14 using point mutation PCR. The digestion of the plasmid with *Eco*RI and subsequent *in vitro* transcription will yield a RNA that ends after the D1 sequence. Since the 3'-overhang is now no longer present in *Sc*.D1-L1 for binding of T-biotin and therefore, biotin labeling, the *Sc*.D1-L1 was postranscriptionally and covalently biotinylated at the 3'-end (Figure 62b).

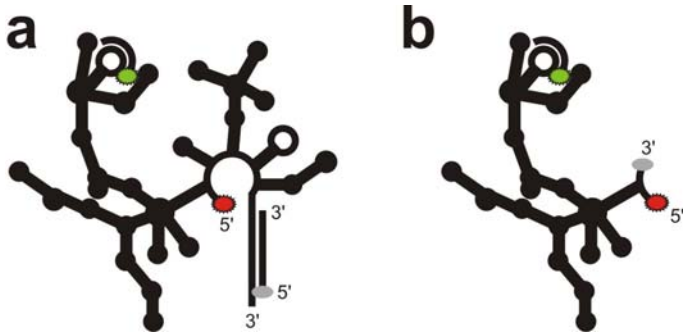


Figure 62. Schematic secondary structure depicting of **a** *Sc*.D135-L14-Cy5 bound to Cy3-DNA and T-Biotin, **b** *Sc*.D1-L1 bound to Cy3-DNA. Cy5 and Biotin are covalently attached to *Sc*.D1-L1 at the 5'- and the 3'-end, respectively. Cy3 (green) and Cy5 (red) are depicted as spikes and biotin is shown as gray marker.

The *in vitro* transcriptions of both *Sc*.D1-L1-GMPS and *Sc*.D135-L14-GMPS were carried out with a mixture of 4:1 GMPS:GTP. Therefore, 80 % of the generated RNAs should

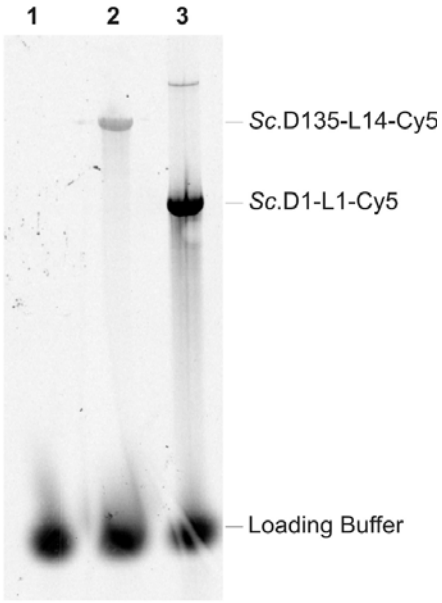


Figure 63. Fluorescence detection of *Sc*.D135-L14-Cy5 and *Sc*.D1-L1-Cy5 on a denaturing gel. Successful labeling of *Sc*.D135-L14-GMPS and *Sc*.D1-L1-GMPS (lane 3) with Cy5 maleimide is proven by detection of Cy5-fluorescence bands for *Sc*.D135-L14-Cy5 (lane 2) and *Sc*.D1-L1-Cy5 (lane 3). A minor side band is detected for *Sc*.D1-L1-Cy5 (lane 3) and might correspond to some form of dimer.

start with a 5'-monothiophosphate. Surprisingly, the transcription yields were hardly reduced and similar to the yields obtained from transcriptions in the absence of GMPS. Yet, subsequent attempts to label the GMPS-carrying RNAs with Cy5-maleimide were partially successful, but so far unsatisfactory. Even repeated labeling of already partially labeled RNAs under optimized conditions would not result in labeling efficiencies of more than maximally 14% for *Sc*.D1-L1-GMPS or 2.3% for *Sc*.D135-L14-GMPS. But a minimal amount of labeled RNA can clearly be detected via UV-absorption of Cy5 (data not shown) or via denaturing PAGE separation and detection of Cy5-emission (Figure 63). The unsatisfactory labeling efficiency is most probably due to the large size (418 nts for *Sc*.D1-L1 and 638 nts for *Sc*.D135-L14) and possibly some form of secondary and tertiary structure of the RNAs during the labeling reaction that complicates the accessibility of the 5'-end and hampers the labeling reaction. The labeling of large RNAs has been described to be increasingly difficult with size.[235] In addition, it seems likely that less than 80% of the RNAs are primed with GMPS when comparing the surprisingly similar transcription yields compared to transcriptions in the absence of GMPS, which might be due to the presence of more GTP in the transcription than anticipated.

Initial smFRET experiments on *Sc*.D1-L1-Cy5 already revealed complications generated by the low labeling efficiency (14 %) and the fact, that the RNA is most probably not 100 % biotinylated and maybe also partially suffered from the harsh conditions during the oxidation and subsequent biotinylation process. Therefore, the immobilization efficiency is reduced and, once immobilized, only 14 % of the molecules carry a donor and an acceptor molecule at the same time. The RNA concentration in the samples needed to be elevated for normal single molecule dot density on the quartz slide and as a result, aggregations and therefore, low resolution hampered attempts to carry out meaningful smFRET experiments.

4.4.2 Acceptor molecule at 3'-end

As an alternative to the 5'-end positioning, the Cy5-label was attached to the 3'-end. The ribozymes are immobilized to the quartz surface in smFRET experiments via binding of a biotinylated oligonucleotide at their 3'-end. This DNA-oligo carries the biotin moiety at its 5'-end but contains a 3'-end that is available for fluorophore labeling. Therefore, the 3'-Cy5-labeled DNA-oligo T-Biotin-Cy5 was generated to bind and immobilize *Sc*.D135-L14 and at the same time introduce the acceptor label to the respective 3'-end (Figure 64).

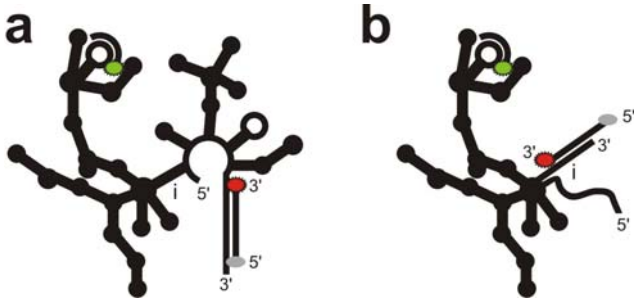


Figure 64. SmFRET set-up for constructs carrying the Cy5-label at the 3'-end. **a** *Sc*.D135-L14 bound to Cy3-DNA and biotin- as well as Cy5-labeled T-Biotin-Cy5. **b** *Sc*.D1-L1 bound to Cy3-DNA and corresponding DNA-oligo T-D1-Cy5 showing disruption of the i-arm in D1. Cy3 (green) and Cy5 (red) are depicted as spikes and biotin is shown as a gray marker.

Initial smFRET experiments with *Sc*.D135-L14 bound to Cy3-DNA and T-Biotin-Cy5 were conducted (Figure 64a). At 10 mM Mg^{2+}, the molecule predominantly stay in a low FRET state at a ratio of ~0.15 (Figure 65a). As found in smFRET with *Sc*.D135-L14*, ~50 % of the molecules show dynamics and higher FRET values in the presence of 100 mM Mg^{2+} (Figure 65b). Again, the molecules are able to adapt three distinct conformational states at different FRET values (Figure 65c). Yet, the low FRET value is no longer at 0.25, but lowered to ~0.15. Such a low ratio is increasingly difficult to be distinguished from photobleached states and hence, this system is not optimal for such studies. Yet, the increase in dynamics at high Mg^{2+} concentrations and the existence of three conformational states is confirmed.

We also attempted to design a similar system to study D1 in isolation. A DNA-oligo is needed to immobilize *Sc*.D1-L1 on the quartz surface. This oligonucleotide, called T-D1-Cy5, would be complementary to the last 20-22 nts of the *Sc*.D1-L1 sequence and carry a biotin at is 5'- and the Cy5-label on its 3'-end. Possible complication can arise due to the partial competition of formation of the i-arm in D1 and the binding of T-D1-Cy5 to the same sequence. Initial attempts to work with this system were unsuccessful, because the T-D1-Cy5 sequence was unluckily designed to bind to the 3'-half of D2, a sequence that was no longer contained in *Sc*.D1-L1. Therefore, the ribozymes would not bind T-D1-Cy5 and subsequently, no be immobilized on the quartz slide and no RNA signals are detected.

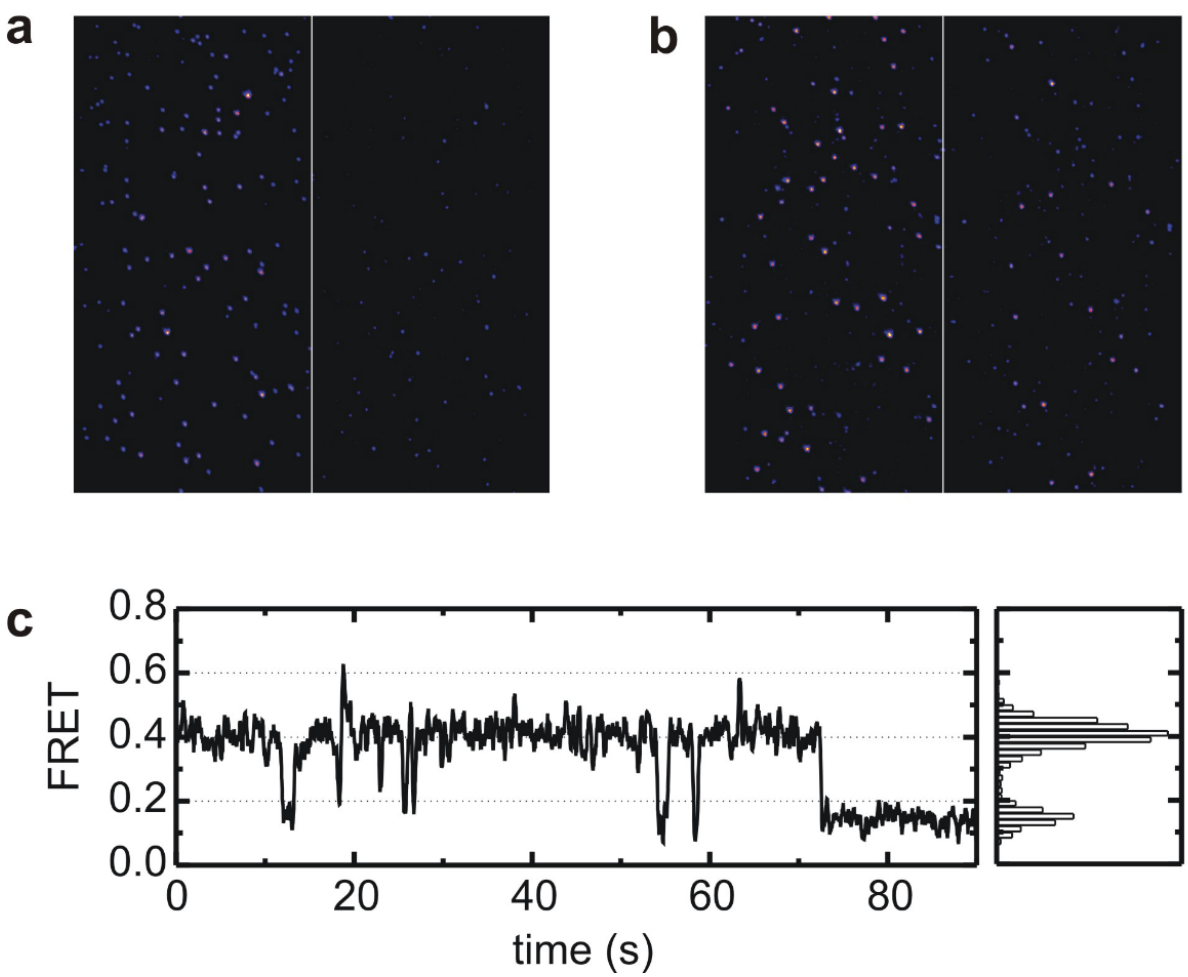


Figure 65. Single molecule experiments using T-Biotin2 (5'-Cy5 labeled). **a** Single molecule spots at 10 mM Mg^{2+} detected in the donor channel (112 spots) are shown to the left, spots detected in the acceptor channel (6 spots) to the right. **b** Single molecule snapshot at 100 mM Mg^{2+} showing 82 spots in the donor (left) and 46 spots in the acceptor channel (right). Numbers of detected spots stated in acceptor channel are spots with average FRET ratio intensities > 0.25. **c** Single molecule trace at 100 mM Mg^{2+} showing three reoccurring FRET states at mean FRET ratios of 0.15, 0.4 and ~0.6.

Further attempts to place the Cy5-analog Alexa Fluor 647 directly to the 3'-end of *Sc*.D1-L1 and also a shorter RNA D1ssh (D1 supershort, 45 nts) were so far unsuccessful and yielded no fluorescently labeled products. The 3'-end of the RNAs were oxidized to carboxylic acids and were then tried to react with the carboxylic acid-reactive Alexa Fluor 647 cadaverine (an aliphatic amine). This reaction is found to be highly delicate and the reactivity drops with decreasing pH.[142]

4.4.3 The constructs *Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ

The original labeling approach using two modular hairpin loops remains the most flexible and promising one. Three more constructs were designed to study conformational changes within D1 during the folding pathway in greater detail: *Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ (Figure 66). The location of the loop in d2b of D1 (L1) that binds Cy3-DNA is

retained in all constructs, but the position of the second loop (L2) was moved throughout the ribozyme molecules. In *Sc*.D135-L1α, L2 replaces nucleotides 36-45 of *Sc*.D135 and hence, L2 is adjacent to the stem-loop comprising the nucleotides termed α that are involved in the tertiary contact α–α' (Figure 66a). L2 is positioned at the original nucleotides 159-161 in *Sc*.D135-L1β. L2 is thereby close to nucleotides involved in the β-part of the tertiary interaction β–β' and hence, the name *Sc*.D135-L1β (Figure 66b). The third construct *Sc*.D135-L1κ harbors L2 right next to the bulge nucleotides, which are known to make up parts of the κ–κ' interaction (Figure 66c).

Both *Sc*.D135-L1α and *Sc*.D135-L1β are designed to study the intramolecular long range tertiary structure formations α–α' and β–β' within D1 and therefore, to report on the slow step of D1 folding as experienced by the selected (labeled) positions relative to each other. The third construct *Sc*.D135-L1κ would not report directly on the κ–κ' tertiary contact formation, since this is an interdomain contact between D1 and D5 whereas both L1 and L2 sit on D1 in *Sc*.D135-L1κ, but rather reflect the specific collapse within D1[56] upon metal ion binding to the κ-ζ element and potentially allow further insights in the proposed switch-mechanism that is accompanied by metal ion binding.[225]

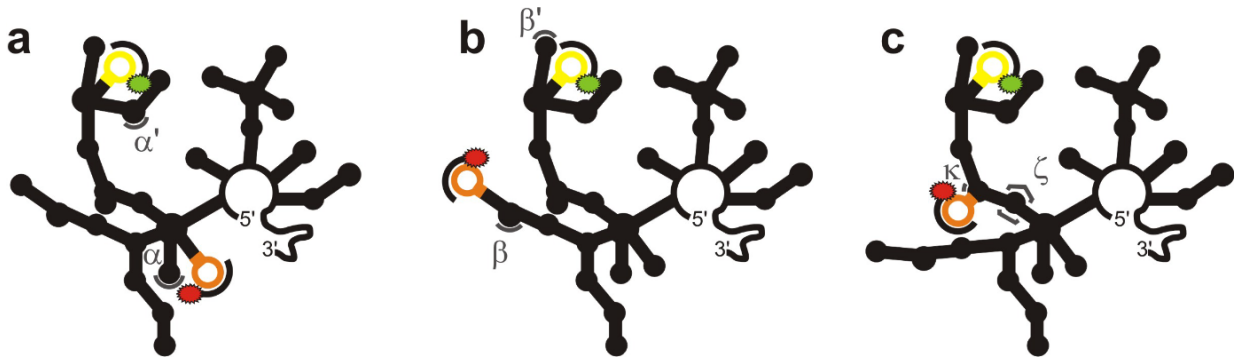


Figure 66. Secondary structure depiction of **a** *Sc*.D135-L1α, **b** *Sc*.D135-L1β and **c** *Sc*.D135-L1κ bound to Cy5-DNA and Cy3-DNA. Cy3 and Cy5 are indicated as green or red spike, respectively. Position of loop 1 (L1) that binds Cy3-DNA is shown in yellow, L2, to which Cy5-DNA binds, is shown in orange. Locations of nucleotides involved in formation of the tertiary contacts α-α', β-β' and the κ-ζ region are marked in grey.

Fluorophore labeled DNA-oligo binding

Native gel analysis was carried out to test for the accessibility of L1 and L2 and binding of Cy3-DNA and Cy5-DNA to the loops on the new constructs. As found for *Sc*.D135-L14*, a part of the fluorophore bound RNA complex sticks in the wells and is not mobile, whereas there are distinct bands that move into the gel (Figure 67). The band that moves furthest into the gel shows a colocalization (but not perfect overlay) of Cy3-DNA and Cy5-DNA for *Sc*.D135-L1α.(Figure 67, lane 1). The band at the same height for *Sc*.D135-L1β is less intense

and a second band appears running slightly higher (Figure 67, lane 2). A color separation seems to occur as the lower band is predominantly green (Cy3-DNA bound) and the upper band appears mainly red (Cy5-DNA bound). In the case of *Sc*.D135-L1κ, it is this second lowest band that predominantly colocalized Cy3-DNA and Cy5-DNA (Figure 67, lane 3).

It is not possible to assign conformational states to the bands on the bases of this native gel only, but it seems that the most compact Cy3-DNA and Cy5-DNA bound state of the ribozymes that moves furthest into the gel is detected for *Sc*.D135-L1α and the same state is still detected for *Sc*.D135-L1β. Yet formation of this most compact state is reduced in the case of *Sc*.D135-L1κ indicating possible interference with sequence modification and/or FRET labeling on the folding pathway.

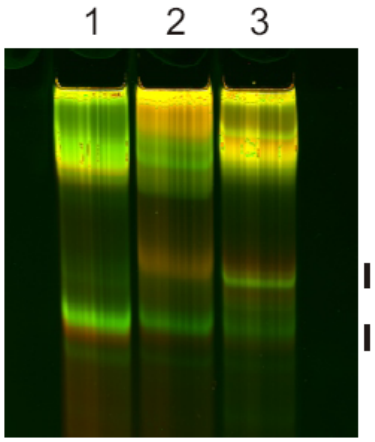


Figure 67. Binding of Cy3-DNA and Cy5-DNA to the additional smFRET constructs detected on non-denaturing gel electrophoresis. Fluorescence detection of bands containing Cy3 (green) and Cy5 (red) to **a** *Sc*.D135-L1α, **b** *Sc*.D135-L1β, **c** *Sc*.D135-L1κ. The two lowest RNA-DNA hybrid bands are indicated as black bars to the right. Overlaying bands showing Cy3 as well as Cy5 emission and appear yellow. Amount of Cy3-DNA and Cy5-DNA is 10 nmol and 20 nmol, respectively, RNA is present at 15 nmol. All samples are prefolded in metal ions (0.5 M KCl, 6 mM $MgCl_2$) before native gel (6% w/v) electrophoresis at 4 °C (3 mM MgOAc and 66 mM HEPES, 34 mM Tris, pH 7.4 (RT))

Catalytic performance

Catalytic assays to test for cleavage of 5'-^{32}P-17/7 result in k_{obs} of 0.30 ± 0.02 min^{-1}, 0.35 ± 0.02 min^{-1} and 0.23 ± 0.02 min^{-1} for *Sc*.D135-L1α*, *Sc*.D135-L1β* and *Sc*.D135-L1κ* (all fully Cy3-DNA, Cy5-DNA and T-Biotin bound), respectively (Figure 68). The product formation was 85% for *Sc*.D135-L1α*, *Sc*.D135-L1β* and a little reduced to 78% for *Sc*.D135-L1κ*. *Sc*.D135-L1α* as well as *Sc*.D135-L1β* cleave slightly faster than *Sc*.D135-L14* (k_{obs} = 0.27 ± 0.01 min^{-1}, Table 7),

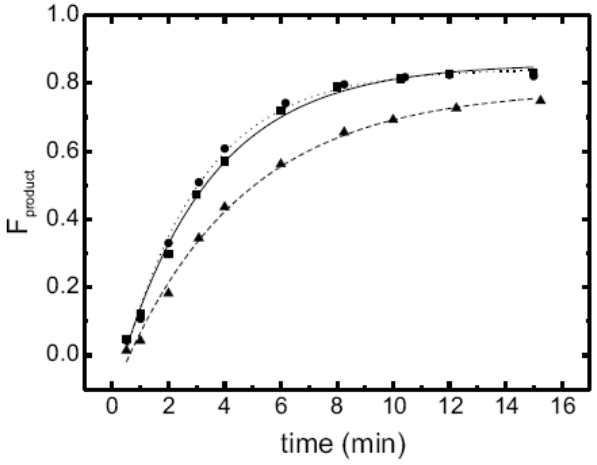


Figure 68. Cleavage of substrate 17/7 (5'-^{32}P) by additional smFRET constructs. Cleavage by *Sc*.D135-L1α* (■, solid line), *Sc*.D135-L1β* (●, dotted line) and *Sc*.D135-L1κ* (▲, dashed line) The lines correspond to single exponential fits and reactions were performed in 80 mM MOPS, pH 7.0, 500 mM KCl, 100 mM $MgCl_2$ and at 42°C.

whereas the cleavage by *Sc*.D135-L1κ* is slowed down.

Hence, *Sc*.D135-L1α* and *Sc*.D135-L1β* are suited to investigate D1 folding by FRET. The *Sc*.D135-L1κ* ribozyme is slightly less active. This is reflected in the reduced cleavage rate and the extent of product formation, although both values are not significantly different from parameters for *Sc*.D135-L14*. An inhibition of catalysis is actually expected in the latter case, because the modular loop is placed at the core of the ribozyme. Taking this in account, the almost equal catalytic performance of *Sc*.D135-L14* is actually surprising.

Overall ensemble folding

The overall FRET distance change upon folding of *Sc*.D135-L1α*, *Sc*.D135-L1β* and *Sc*.D135-L1κ* was characterized in steady-state FRET measurements analogous to the experiments with *Sc*.D135-L14* (Figure 47).

Surprisingly, the FRET ratio in *Sc*.D135-L1α* does not increase upon Mg^{2+} addition up to 100 mM (Figure 69a)! Rather, the FRET value is decreasing in an approximately linear fashion from 0.28 to 0.265. The folding rates do not change in dependency of Mg^{2+} and at 100 mM, k_{obs} = 1.2 ± 0.1

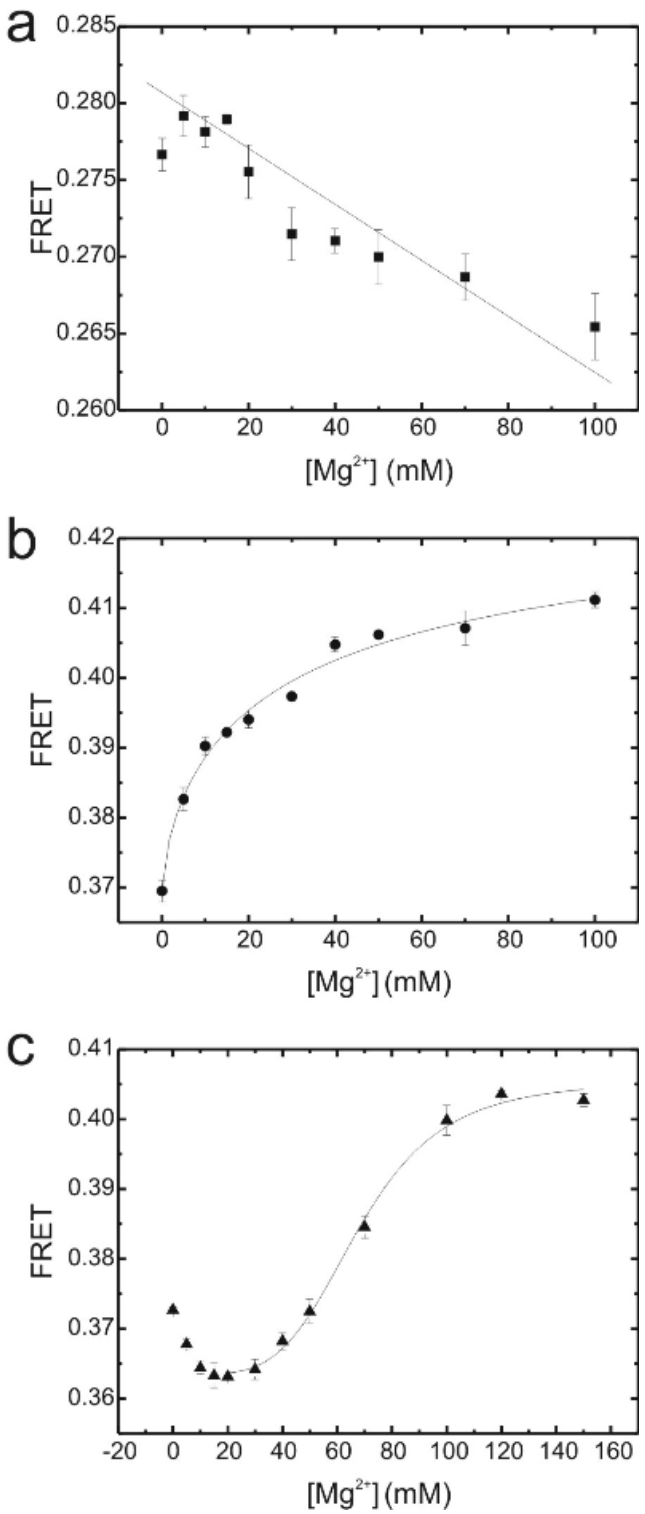


Figure 69. FRET ratios of additional smFRET constructs dependent on Mg^{2+} concentration. **a** The FRET ratio of *Sc*.D135-L1α* follows a linear decrease upon increasing amounts of Mg^{2+} (0-100 mM). **b** FRET ratio of *Sc*.D135-L1β* increases with a dependency of K_{Mg} = 30 ± 23 mM upon Mg^{2+} addition. Solid black line depicts fit with Hill equation. **c** *Sc*.D135-L1κ* shows a decreasing FRET ratio in the range of 0-20 Mg^{2+} before value increases to a maximum. The increase during 20-150 mM Mg^{2+} is fitted with the Hill equation (solid black line) resulting in a K_{Mg} = 68 ± 3 mM. Samples are recorded in 80 mM MOPS, pH 7.0, 500 mM KCl, $MgCl_2$ as indicated and at 42 °C.

min^{-1}. It appears that the fluorophores at their positions on *Sc*.D135-L1α* do not undergo large relative distance changes upon folding. Rather, the two selected regions diverge even more upon folding and do not report on the formation of the tertiary contact α–α'. These results make *Sc*.D135-L1α* impracticable to study formation of α–α' by FRET. Apart from this, an overall folding rate of k_{obs} ~ 1 min^{-1} of D1 compaction is, again, confirmed.

In the case of *Sc*.D135-L1β*, the overall FRET ratio is higher than observed for *Sc*.D135-L1α* and increases from 0.37 to 0.43 upon addition of 100 mM Mg^{2+} (Figure 69b). A dependency of the final FRET value on Mg^{2+} concentration is recorded and can even be fitted with a Hill equation resulting in K_{Mg} = 30 ± 23 mM. As the very large error indicates, this fit does not describe the observed data well in any way. The rate of folding again does not change much with Mg^{2+} and stays approximately equal at k_{obs} = 1.5 ± 0.2 min^{-1} (100 mM Mg^{2+}). This rate indicates a slightly faster folding step within D1 and might reflect the formation of the long-range β–β' tertiary contact formation.

The observed overall FRET ratio of *Sc*.D135-L1κ* first slightly decreases from 0.373 to 0.363 as the Mg^{2+} concentration is increased from 0 to 20 mM and then increases up to 0.405 by further Mg^{2+} addition up to 150 mM (Figure 69c). This result is very interesting! It appears that two separate events can be resolved in construct *Sc*.D135-L1κ*. The initial decrease from 0-20 mM Mg^{2+} corresponds to the concentration range, in which the initial collapse within D1 occurs.[132] The FRET decrease in *Sc*.D135-L1κ* at 0-20 mM might therefore report on the binding of a metal ion to the κ-ζ element that results in the formation of I.[56, 225] The increase in FRET ratio at concentrations of 20-150 mM Mg^{2+} is described with a Mg^{2+} dependency of K_{Mg} = 68 ± 3 mM. The folding rate is again found not to change dramatically with Mg^{2+} and shows a k_{obs} = 0.55 ± 0.10 min^{-1} at 100 mM. The high K_{Mg} and the reduced k_{obs} of *Sc*.D135-L1κ* of the second event clearly reveal deviations from the Mg^{2+} requirements and the overall folding pathway of *Sc*.D135 and *Sc*.D135-L14*. *Sc*.D135-L1κ* shows an altered behavior in the formation of (near-) native states (F and N), but reports on two distinct events that are comfortably separated to be individually studied.

Taken together, the most promising construct for further smFRET experiments are *Sc*.D135-L1β*, and, to a lesser extent, *Sc*.D135-L1α*. *Sc*.D135-L1β* proves to be fluorophore bound, active and experiences a distinct change in FRET upon Mg^{2+} addition. Furthermore, the folding rate (k_{obs} ~ 1.5 min^{-1}) of *Sc*.D135-L1β* is 50 % faster than the rate of folding detected by Sc.D135-L14*(k_{obs} ~ 1 min^{-1}). This result indicates a conformational change at an early stage of the folding pathway that potentially corresponds to the formation

of the long-range tertiary contact β-β'. *Sc*.D135-L1κ* is a very interesting construct, too. Although *Sc*.D135-L1κ* needs very high amounts of Mg^{2+} to reach the most compact state, it is still able to do so. In addition, it is the only construct that is able to distinguish between two events that possibly correspond to formation of I from U followed by folding to F/N.

4.5 Final remarks and Outlook

Sc.D135-L1α* is the construct, in which the sequence modification and the labeling scheme overall apparently interferes the least with the folding and cleavage reaction. Unluckily, the regions chosen to carry the fluorophore labels do not seem to undergo large distance changes upon overall folding. Nonetheless, it would be interesting to perform smFRET measurements on *Sc*.D135-L1α* at least in a trial experiment, since one never knows what the single molecule technique is able to reveal!

Sc.D135-L1κ* on the other hand shows deviations from the *Sc*.D135 system. The sequence modifications and the binding of the DNA-oligos effects catalysis only minimally, yet the Mg^{2+} requirements for folding to the most compact state are increased compared to *Sc*.D135-L14*. The κ-ζ region is known to be of elementary importance and hence, is of primary interest for the folding pathway. It is not a surprise that an RNA that is considerably modified and enlarged very close to this folding control element is affected in folding. Yet, it is interesting to note that high amount of Mg^{2+} are able to overcome this folding problem and turn the *Sc*.D135-L1κ* into a considerably catalytically competent ribozyme. In addition, the observed decrease in FRET ratio between 0-20 mM Mg^{2+} might correspond to the capturing of the intermediate conformational state (I). It will be very interesting to learn more about *Sc*.D135-L1κ* in smFRET experiments.

All posttranscriptional attempts to modify this very large ribozyme system have not yet been fully successful. Nevertheless, the 3'-labeling approach described in the Section 4.4.2 is found to be applicable and could be an option to study folding of D1 in isolation. Rather than having the Cy5-label near the 3'-end of D1, isolated D1 could also be studied on constructs using the loops that bind Cy5-DNA as described in *Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ) yet are missing D2-D5. Such constructs (*Sc*.D1-L1α, *Sc*.D1-L1β, *Sc*.D1-L1κ) can be generated as shown for *Sc*.D1-L1.

Further smFRET experiments with the here presented or additional constructs can be conducted in the presence of substrate(s) and different metal ion conditions or any other type of cofactor that can be thought of. Initial examples are given in sections 3.2.11 and 3.2.12. Importantly, experiments in the presence of a protein cofactor, i.e. Mss116, will be studied in the near future to deduce the significance of protein binding to group II intron ribozyme folding and to contribute to up-to-date research and discussion in this field.

In an immediate next step, it will be important to carry out smFRET experiments at temperatures other than room temperatures to test the newly stated paradigm for folding of

large RNAs under different conditions. These experiments will become possible once the temperature control extension has been added to the smFRET set-up.

The modular character of the RNA system as well as the smFRET technique allows a diverse array of experiments. Depending on the focus, specific experiments can be designed to collect pieces of information on RNA systems that are even as large as the *Sc*.D135-L14* ribozyme complex.

5. Materials and Methods

5.1 Materials and Chemicals

DNA oligonucleotides were purchased from Microsynth, Balgach (Switzerland) or from HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory, Yale University, New Haven CT (USA). The nucleotide triphosphate came from Amersham Biosciences, Otelfingen (Switzerland), now part of GE Healthcare, Glattbrugg (Switzerland), except for UTP, which was obtained from Sigma-Aldrich-Fluka, Buchs (Switzerland), 2-NH_2-PuMP was purchase from BioLog, Bremen (Germany). RNA was either transcribed *in vitro* or purchased from HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory, Yale University, New Haven CT (USA) or Dharmacon Inc., Lafayette, CO (USA).

Polynucleotide Kinase (PNK), *Pfu* polymerase and restriction enzymes *Hind*III, *Eco*RI and *Dpn*I were from Promega, Madison WI (USA). ^{32}P labeled nucleotide triphosphates and Cy5-maleimide were purchased from Amersham Biosciences, Otelfingen (Switzerland), now part of GE Healthcare, Glattbrugg (Switzerland). Alexa Fluor 647 cadaverine was purchased from Invitrogen AG, Basel (Switzerland) and T7 polymerase used for *in vitro* transcription was homemade.[239, 271]

PCR products were purified using Perfectprep Gel Cleanup Kit from Eppendorf AG, Hamburg (Germany). Plasmid minipreps were isolated using High Pure Plasmid Isolation Kit from Roche, Basel (Switzerland) and plasmid maxipreps were purified using Plasmid Maxi Kit from BioRad Laboratories, Hercules CA (USA) or HiSpeed Plasmid Maxi Kit from Qiagen, Hilden (Germany). Digested plasmids were purified with Plasmid Midi Kit from Qiagen, Hilden (Germany) or using the Wizard® Clean-up System from Promega, Madison WI (USA). Sephadex MicroSpin G50 size exclusion columns for desalting and removal of unicorporated nucleotides were purchased from Amersham Biosciences, Otelfingen (Switzerland), now part of GE Healthcare, Glattbrugg (Switzerland).

The $MgCl_2$ used in the metal ion titrations was obtained as 1 M ultrapure solution in H_2O from Fluka. The exact concentrations of the $MgCl_2$, $CaCl_2$, $MnCl_2$ and $CdCl_2$ stock solutions in 99.999% D_2O each (Sigma-Aldrich) was determined by potentiometric pH titration employing EDTA. $TbCl_3$ came from Sigma-Aldrich.

All chemicals used were at least puriss p.a. and purchased from either Fluka-Sigma-Aldrich or Brunschwig Chemie, Amsterdam (The Netherlands).

5.2 Instrumentation

The *Biotrap* apparatus used for electroelution was purchased from Schleicher & Schuell, Dassel (Germany). Centrifuges were a 5415R and a5804R with rotor F-45-24-11 and 5804R with rotor A-4-44 from Eppendorf, Hamburg (Germany), as well as a RC5C plus from Sorvall, Langenselbold (Germany) with a SA-600, a SH-3000 and a SLA-3000 rotor. The purified RNA samples were vacuum-dried in a Concentrator 5301 from Eppendorf, Hamburg (Germany). Gels were dried on a Biometra Maxidry geldryer, Goettingen (Germany).

The HPLC system used for RNA purification is a Merck Hitachi Elite La Chrom. The column used for ion exchange chromatography is an analytical 4x250 mm DNAPac® PA-100 from Dionex, Sunnyvale CA (USA). For reversed phase chromatography a C18 Beckmann Coulter Ultrasphere column was employed.

UV measurements were carried out on a Varian Cary 500 Scan UV-VIS-NIR spectrophotometer connected to a circulating temperature controller using a 1 or 10 mm QS cuvette from Hellma, Müllheim (Germany).

2AP fluorescence experiments were performed on a Perkin-Elmer luminescence spectrometer LB-50 connected to a circulating waterbath. Steady-state and anisotropy FRET experiments were recorded on a Cary Eclipse fluorescence spectrometer from Varian Inc., Palo Alto CA (USA) and a Cary PCB 150 water peltier system using a 3 mm square cuvette from Hellma, Müllheim (Germany). Native FRET gels were imaged on a Molecular Dynamic Typhoon 9400 imager from Amersham Biosciences, Otelfingen (Switzerland), now part of GE Healthcare.

Scintillation counts were performed on a scintillation counter 22000CA Liquid Scintillation Analyzer from Canberra Packard. Phosphoimaging was recorded on a Storm 860 phosphoimager from Amersham Biosciences, Otelfingen (Switzerland), now part of GE Healthcare.

The H_2O used in all experiments was treated with a TKA genepure water purification system from TKA Wasseraufbereitungssysteme, Niederelbert (Germany).

SmFRET and trFRET experiments were carried out using the equipment in the lab of Prof. D. Rueda at Wayne State University:[272] Single molecule FRET experiments were recorded on a total internal reflection microscope, where fluorescence excitation originated from total internal reflection of a circularly polarized 514 nm laser beam from an argon laser off the quartz surface. The microscope device is composed of a water immersion objective with a 1.2 numerical aperture, mounted on an inverted microscope. The emission was

recorded with a 12-bit cooled, intensified CCD (charged coupled device) camera. A 540-nm long-pass filter was used to block laser light scattered from defects in the quartz and the fluorescence emission was split with a dichroic 635 nm long-pass beamsplitter. The buffer exchange experiments were conducted using a home-built automated buffer exchange system. For trFRET experiments, the donor dye was repetitively excited at 514 nm with 90 ps-duration pulses from a mode-locked argon ion laser. Isotropic fluorescence decay was detected up to at least 40'000 single-photon counts under magic angle conditions at 580 nm using a microchannel plate photomultiplier. The output was processed with a time-correlated single-photon counting system. The instrument response function (fxhm = 100 ps) was measured using a dilute solution of nondairy coffee creamer to scatter the laser pulses.

5.3 Methods Chapter 2 (*Av*.D135)

5.3.1 RNA preparation

100 µg plasmid pEL08, which contains the *Av*.D135 sequence, was linearized over night with *Hind*III and subsequently purified to result in 50.8 µg cut plasmid (50.8 % yield). The RNA was transcribed in a 5 mL *in vitro* transcription assay using 5 µg restriction digested plasmid pEL08, 1 mM ATP, CTP; GTP, and UTP in 1x Transcription Buffer (40 mM TrisHCl, pH 7.5, 30 mM $MgCl_2$, 40 mM DTT, 10 mM Spermidine, 0.01 % Triton X-100) and 10 µL/mL selfmade T7 RNA polymerase.[271] The reaction was allowed to proceed for 3.25 h at 37 °C and shaking (300 rpm). Transcription products were precipitated by adding 250 mM NaCl and 3 volumes of ethanol and purified by denaturing 5% (w/v) PAGE. The RNA band was located by UV-shadowing, excised and eluted in Elution Buffer (10 mM MOPS, pH 6.0, 1 mM EDTA, 250 mM NaCl) at 4 °C for 2 h. Eluted RNA was precipitated with 3 volumes of ethanol and RNA pellets were dissolved in Storage Buffer (10 mM MOPS pH 6.25, 10 µM EDTA). The *Av*.D135 concentration was determined spectroscopically from absorbance at 260 nm using the equation for concentration of large RNAs (c = A/(#NTP/100), #NTP = 628) yielding 0.73 µmol/mL *Av*.D135.

5.3.2 RNA labeling

Av.D135 was 3'-labeled with α-P^{32}-dCTP using the Szostak/Splint method.[273] In a total volume of 60 µL, 6 µM *Av*.D135 was heated to 95 °C for 30 s together with 12 µM oligo

splint (5'- CGGAGCTTGCATG-3') in Annealing Buffer (14 mM TrisHCl, pH 7.5, 40 mM NaCl, 0.2 mM EDTA), then the mixture was left to cool to 42 °C over 1 h for annealing. 45 µL α-^{32}P-dCTP (10 µCi/µL) was added to the annealing mixture and labeling reaction performed using Klenow DNA Polymerase for 2 h at 37 °C and shacking (300 rpm) in 7 mM $MgCl_2$, 1 mM DTT.

Unincorporated nucleotides and salt were removed by passing the mixture over G50 columns. The resulting flowthrough was purified on a denaturing 5% PAGE and full-length *Av*.D135 was located radiographically. Radioactively labeled *Av*.D135 was excised and eluted into Elution Buffer for 1 h at 4 °C, precipitated in 3 volumes of ethanol and dissolved in Storage Buffer.

5.3.3 Tb^{3+} stock solutions

$TbCl_3$ stock solutions were prepared in a buffer of 5 mM Na-cacodylate, pH 5.5, to prevent formation of insoluble hydroxide species.

5.3.4 Tb^{3+} cleavage reaction

In a series of aliquots of a total volume of 5 µL, trace amounts (~1 nM) of 3'-labeled *Av*.D135 were combined with 1 µM unlabeled *Av*.D135 and heat-annealed in Tb^{3+}-cleavage buffer (25 mM MOPS, pH 7.0, 0.5 M KCl) for 45 s at 95°C. The samples were immediately transferred to 42 °C and $MgCl_2$ was added to a total concentration of 100 mM each. Then, the mixtures were incubated at 42 °C for 10 min. $TbCl_3$ was added to final concentrations of 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.75, 1, 2.5, 5, 7.5, 10 and 25 mM. Cleavage reactions were performed over 1h on ice before 5 µL formamide loading buffer (FLB: 80% (v/v) formamide, 0.025% (w/v) xylene cyanol, 0.025% (w/v) bromphenol blue, 50 mM EDTA) was added to stop reaction. Samples were precipitated with 3 volumes of ethanol, then resuspended in 6 µL FLB and analyzed on 18% and 5% PAGE.

An alkaline hydrolysis cleavage ladder was generated by incubating *Av*.D135 (trace amount of labeled RNA in a background of 1 µM unlabeled RNA as stated above) in 200 mM $NaHCO_3$, pH 8.0 at 75 °C for 5 min. The cleavage reaction was quenched by addition of FLB. The T1-ladder, a G-specific sequencing ladder from partial digestion of the RNA with ribonuclease T1 (cuts after G), was generated by digesting *Av*.D135 with tRNA (0.04 mg) as a background and RNase T1 (100 mU) in 50 mM TRIS/HCl pH 7.5, 2 mM EDTA at 37 °C for

15 min. The reaction was stopped by addition of FLB. Both samples were loaded onto gels for analysis.

5.3.5 Product band classification

Bands on 15% polyacrylamide gel (Figure 22) were quantified using ImageQuant Software Version 5.2 from Molecular Dynamics (www.mdyn.com). Since single nucleotides could not be resolved over the entire gel because bands overlay towards the top of the gel due to closer spacing of nucleotide bands, quantification was difficult and product bands were classified at 2.5 mM Tb^{3+} into weak (5'000-13'000 cpm) and distinct (13'000-30'000 cpm) binding sites only.

5.4. Methods Chapter 3 (D5-36-2AP)

5.4.1 Preparation of RNA oligonucleotides

Oligonucleotides containing 2AP

The 2AP labeled RNAs D5-19-2AP (5'-AAGUCGC(2AP)CGUACGGUUCC-3'), D5-36-2AP (5'-GGAGCCGUAUGCGAUGAAAGUCGC(2AP)CGUACGGUUCC-3') and D5-38-2AP (5'-GGAGCCGUACGUGCGAUGAAAGUCGC(2AP)CGUACGGUUCC-3') were purchased from Dharmacon, Inc. (Lafayette, CO), dissolved in Storage Buffer (10 mM MOPS pH 6, 10 μM EDTA) and checked for purity by HPLC using the ion exchange column. The binding buffer was 0.02 AcOH/AcoNa, pH 6.0, 20% MeCN and RNA was eluted with a gradient from 0-1 M KCl over 50 min at a flow of 1 mL/min. Samples stored for more than one year were HPLC-purified on a reversed phase chromatography column preequilibrated in 0.1 M triethylammonium acetate (TEAA) at pH 7.0 and eluted using an optimized gradient of 7.5 - 25% acetonitrile (Biosolve) over 45 min at a flow of 1 mL/min according to protocols.[274] The purified RNAs were lyophilized and dissolved in Storage Buffer. The concentrations were determined by UV absorption at 260 nm and the extinction coefficients $\varepsilon_{D5\text{-}19\text{-}2AP}$ = 164.3 $mM^{-1}cm^{-1}$, $\varepsilon_{D5\text{-}36\text{-}2AP}$ = 328.5 $mM^{-1}cm^{-1}$ and $\varepsilon_{D5\text{-}38\text{-}2AP}$ = 346.0 $mM^{-1}cm^{-1}$ (http://eu.idtdna.com/analyzer/Applications/ OligoAnalyzer).

Unlabeled oligonucleotides and 2-NH_2-PuMP

D5-36 (5'-GGAGCCGUAUGCGAUGAAAGUCGCACGUACGGUUCC-3') and D5-17 (5'-GGAGCCGUAUGCGAUGA-3') were transcribed by *in vitro* standard transcription

protocols[239] using optimized conditions at 5 mM NTPs, 30 mM $MgCl_2$, 1.2 µM synthetic DNA oligonucleotide templates TS-23 (5'-GAAATTAATACGACTCACTATAG-3') and D5-17-OT (5'-TCATCGCATACGGCTCCTATAGTGAGTCGTATTAATTTC-3') or D5-36-OT (5'- GGAACCGTACGTGCGACTTTCATCGCATACGGCTCCTATAGTGAGTCGTATT-AATTTC-3') for D5-17 and D5-36, respectively. The transcriptions were carried out over night using homemade T7 RNA polymerase[271] in Transcription Buffer (40 mM TrisHCl, pH 7.5, 40 mM DTT, 10 mM Spermidine, 0.01 % Triton X-100). RNA transcripts were purified by denaturing 18% (w/v) polyacrylamide gel elecrophoresis (PAGE) and the RNA band was located by UV shadowing and excised. The RNA was separated from gel by electroelution in 1x TE, precipitated by adding three volumes of ethanol and dissolved in Storage Buffer. Transcription yields were 0.04 mg/mL and 0.24 mg/mL for D5-17 and D5-36, respectively.

2-NH_2-PuMP was dissolved without further purification in Storage Buffer.

5.4.2 2AP steady-state fluorescence spectroscopy

All experiments were conducted in Standard Buffer (10 mM MOPS, pH 6.0, 10 mM KCl) if not stated otherwise.

Two-piece D5 experiments: D5-17 and D5-19-2AP

D5-17 and D5-19-2AP were mixed in a 1:1 ratio at 10 µM each in a total volume of 100 µL. The solution was heated to 95°C for 1 min and left to cool down to room temperature over 15 min and subsequently transferred into the cuvette. Typically 20 aliquots of either 10 mM $MgCl_2$, $CaCl_2$, $MnCl_2$ or $CdCl_2$ stocks, generated from the 1 M potentiometrically titrated stocks, were sequentially added in 0.5 µL steps to span the final metal ion concentration range of 0-1 mM. After every aliquot addition, the sample was manually mixed by pipeting up and down before the sample was excited at 310 nm and fluorescence emission was recorded from 320 to 420 nm at a scan speed of 100 nm/min. Slits were set to 4 nm and the cuvette was thermostatted at 25°C. Maximal emission at 370 nm of each scan was plotted against corresponding metal ion concentration in graphs for further analysis.

Fluorescence intensities from experiments of two-piece D5 and D5-19-2AP only were normalized via division by the corresponding value in the absence of metal ions. The corrected graphs were obtained by subtraction of the normalized data of D5-19-2AP from the

data of two-piece D5 normalized in the same way from the corresponding experiments with the identical metal ion.

Full-length D5-36-2AP titrations

A 100 µL aliquot of 1 µM full-length D5-36-2AP was heat-annealed at 95 °C for 45 s before solution was incubated for at least 15 min at RT before transferring it into a cuvette thermostatted at 25 °C. 2AP was excited at 310 nm (5 nm slit width) and the emission was recorded from 320-420 nm (15 nm slit width). Addition of divalent metal ions without change in RNA concentration was performed by exchanging 5 µL of the sample in the cuvette by 5 µL of another solution containing 1 µM D5-36-2AP and metal ions in Standard Buffer. The second solutions was prepared the same way as the first solutions apart from the fact that metal ions were added to the solution 1-2 min after the initial heating step to prevent unspecific degradation of the RNA by high metal ion concentrations at high temperatures. The sample was mixed three times by pipeting up and down and the solution was left to reach equilibrium for 4 min before the data collection. Typically, 15 scans were recorded for metal ion concentration up to 15, 30, 62, 125 and 185 µM.

Where present, DTT was also only added after the initial heating step to prevent decomposition of DTT at high temperatures. In experiments with DTT, $MnCl_2$ concentrations were reduced to 0.04 mM in the second solution to optimally record fluorescence change and 15 scans were recorded in the metal ion concentration range of 0-25 µM, whereas $MgCl_2$, $CaCl_2$ and $CdCl_2$ concentration were at 0.1 mM and tested the metal ion concentration range from 0-62 µM.

Measurements with HPLC-purified D5-36-2AP were conducted on a 100 µL aliquot of 1 µM full-length D5-36-2AP heat-annealed as described above. Metal ion titration were performed by addition of 0.5 or 1 µL aliquots of concentrated metal ion stock solutions to give final concentrations of 0, 0.025, 0.05, 0.1, 0.15, 0.2, 0.44, 0.68, 0.91, 1.38, 1.85, 3.7 and 6 mM for Mg^{2+}, Ca^{2+} and 0, 0.015, 0.03, 0.05, 0.08, 0.1, 0.15, 0.2, 0.3, 0.48 and 0.72 for Mn^{2+}. Maximal fluorescence emission at 370 nm was corrected for dilution and plotted versus metal ion concentration in graphs for further analysis.

Control experiments

Control experiments were conducted in the absence of DTT (except in oxygen scavenger experiments) and without correction for dilution. In experiments to test for photobleaching, the samples were either continuously excited at 310 nm while emission was recorded in

parallel at 370 nm or alternatively emission scans from 320-420 nm were recorded in intervals of 6 min. In the experiment to test for fluorescence change by addition of monovalent cations, 15 aliquots (1 µL) of 1.4 M KCl up to a final concentration of 175 mM KCl were sequentially added to the sample in the cuvette. Concentration of 2-NH_2-PuMP in experiments was set to 0.7 µM and slit widths were 4 nm for excitation and emission. Otherwise, the experiments were done analogously to the ones described for the two-piece D5 experiments (see above). The experiments to test for influence of oxygen scavenger were performed in the presence of 25 mM DTT and experiments for Mn^{2+} in the presence and absence of DTT were recorded over a final concentration range of 0-31 µM.

Single site binding equation

The titration curves were fitted with an equation describing the binding of a metal ion to RNA in a 1:1 fashion adapted from [25, 275, 276] multiplied by a term describing a linear decrease to mathematically combine the two separate effects that lead to fluorescence intensity change. The equation can be interpreted as binding of one metal ion to a single site in the RNA:

$$F_{obs} = F_0 + \left(\frac{F_0 - F_{max}}{2 \times R} \times \left(K_D + x + R\right) - \sqrt{\left(K_D + x + R\right)^2 - 4 \times x \times R} \right) \times \left(1 - a \times x\right) \qquad (7)$$

where x is final metal ion concentration, F_0 is intensity at x = 0, F_{max} is maximal intensity, R is RNA concentration, K_D is dissociation constant and a is the slope of linear decrease.

K_D values given in text, figures and tables are averaged numbers from at least three independent measurements, and the curves depicted in graphs are representative examples from individual experiments.

5.4.3 UV melting curves

A series of 250 µL aliquots of 10 µM RNA samples in an aqueous solution of 100 mM KCl were heat-annealed for 1 min at 95°C followed by addition of metal ions to final concentrations of 0, 0.25, 0.5, 1, 2, 4, 6, 8, 10, 16, 20 $MgCl_2$ or 0, 1, 4, 10, 20 mM $CaCl_2$ and incubation for at least 10 min at RT. The samples were centrifuged and degassed for 30 s under vacuum to get rid of dust particles and dissolved oxygen. The samples were transferred to a 0.1 cm cuvette and covered by a layer of 50-100 µL paraffin oil to avoid evaporation. In addition, the cuvette was capped and the cap was fixed by wrapping with parafilm. The

cuvette was then heated with a rate of 0.5 °C/min and the absorption recorded at 260 nm over the temperature range from 20 to 95 °C. The data from three repeated heating and cooling cycles was gathered to check for reproducibility and reversibility and used for analysis. The melting experiments of D5 with $MgCl_2$ were conducted by Siran Lu during a student project preceding this work.

To check for the effect on thermal stability by the substitution of A25 with 2AP in the bulge, melting experiments of D5 and D5-36-2AP in Standard Buffer were measured in 0 mM or 1 mM $MgCl_2$ and otherwise as described above. Data in Figure 35b is normalized to values between 0 (minimum) and 1 (maximum).

Melting curves were fitted with either an equation describing one transition with a linear increasing baseline in the folded state (equation 8) or an equation with two transitions and a linear increasing baseline in the folded state (equation 9). Both equations are derived from van't Hoff analysis as described in [218, 277]:

$$y(x) = \frac{b_f + m_f \times x + A_u \times e^a}{1 + e^a} \tag{8}$$

$$y(x) = \frac{b_f + m_f \times x + A_i \times e^a + A_u \times e^a \times e^b}{1 + e^a + e^a \times e^b} \tag{9}$$

$$a = \Delta H_1 \times \frac{\frac{1}{T_{m_1}} - \frac{1}{x}}{8.31451} \qquad \text{and} \qquad b = \Delta H_2 \times \frac{\frac{1}{T_{m_2}} - \frac{1}{x}}{8.31451}$$

where x is the temperature in °C, b is the absorption at x = 0, m is the melting slope, f stands for folded, u for unfolded, i for intermediate state, A is the absorbance, ΔH is the enthalpy change and T_m is the melting temperature.

5.5 Methods Chapter 4 (smFRET)

5.5.1 Preparation of *Sc*.D135 derived RNAs and DNA-oligonucleotides

***Sc*.D135-L1 and *Sc*.D135-L4.** The *Sc*.D135 ribozyme sequence derived from the *S. cerevisiae* intron *Sc*.ai5γ encoded on the plasmid pT7D135 was modified by insertion of two loops with the length of 18 nts each. The sequence L1 (3'-CCCAAAUUAAUAACGCUCUUGGUAG-GG-5') was placed between residues 276 and 305 on the d2b stem of D1 or the residues 681-806 in D4 (Figure 42) using standard PCR cloning technique giving the plasmids pT7D135-L1 and pT7D135-L1.

The RNAs were transcribed in 5 mL *in vitro* transcriptions using 7.5 μg/mL *Hind*III digested plasmids, 5 mM ATP, CTP, GTP, and UTP in 1x Transcription Buffer (40 mM TrisHCl, pH 7.5, 20 mM $MgCl_2$, 40 mM DTT, 10 mM Spermidine, 0.01 % Triton X-100) and 10 μL/mL selfmade T7 RNA polymerase.[271] The reactions were allowed to proceed for 5-6 h at 37 °C and shaking (300 rpm). Transcription products were precipitated by adding 250 mM NaCl and 3 volumes of ethanol and purified by denaturing 5% (w/v) PAGE. The RNA bands were located by UV-shadowing, excised and eluted in Elution Buffer (10 mM MOPS, pH 6.0, 1 mM EDTA, 250 mM NaCl) at 4 °C for 2 h. Eluted RNA was precipitated with 3 volumes of ethanol and RNA pellets were dissolved in Storage Buffer (10 mM MOPS, pH 6.25, 10 μM EDTA). The RNA concentration was determined spectroscopically from absorbance at 260 nm using the equation for concentration of large RNAs (c = A/(#NTP/100), #NTP = 612 and 634 for *Sc*.D135-L1 and *Sc*.D135-L4, respectively) yielding 0.3 μmol/mL RNA.

***Sc*.D135-L14.** The sequence for *Sc*.D135-L14 was generated by replacing the residues 681-806 in D4 on plasmid pT7D135-L1 by the sequence L2 (3'-CCCAAAUAUAUUU-CGGACUACGUAGGG-5') using PCR (*vide infra*) giving the plasmid pT7D135-L14. Transcription reactions were performed as described for *Sc*.D135-L1 and *Sc*.D135-L4 resulting in equal transcription yields (#NTP = 638) for *Sc*.D135-L14 RNA.

***Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ.** The *Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ RNAs were prepared analogously to *Sc*.D135-L14. The sequence encoded on the plasmid pT7D135-L14 was modified with L2 at positions 36-45, 159-161 or 207-210 for plasmids pT76D135-L1α, pT76D135-L1β or pT76D135-L1κ, respectively. Transcriptions were carried out as described and resulted in roughly equal yields of 0.3 μg/mL for *Sc*.D135-L1α, *Sc*.D135-L1β and *Sc*.D135-L1κ.

***Sc*.D135-L14-Cy5 and *Sc*.D1-L1-Cy5**

The plasmid sequence on pT7D135-L14 was modified by point mutation PCR to harbor an additional *Eco*RI restriction site at the end of D1/D2 to yield the plasmid pT7D1-L1-ER. After the PCR, the product was treated with *Dpn*I for 1 h at 37 °C to digest the original plasmid pT7D135-L14. The plasmid pT7D1-L1-ER was then linearized with *Eco*RI to obtain a construct of D1 only.

Transcription reactions (5 mL) of digested pT7D135-L14 (*Hind*III-digest) and pT7D1-L1-ER (*Eco*RI-digest) were carried out in a mixture of 7.5 µg/mL digested plasmid, 5 mM ATP, CTP and UTP, 1 mM GTP and 4mM GMPS in 1x Transcription Buffer for 5 h at 37 °C using 10 µL/mL selfmade T7 RNA polymerase[271] to yield 5'-GMPS-labeled *Sc*.D1-L1-GMPS and *Sc*.D135-L14-GMPS.

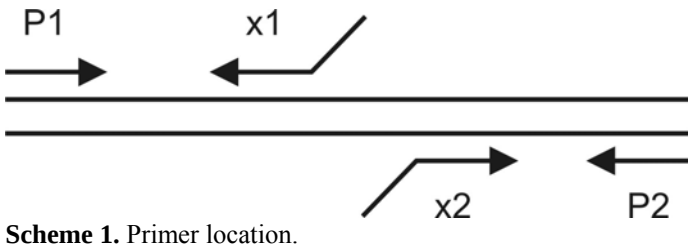
Subsequently, the transcription buffer was exchanged using Vivaspin columns to 10 mM HEPES, pH 7.5, 1 mM $MgCl_2$, the sample volume was reduced to 0.5 mL and any residual molecules (i.e. salts, unincorporated nucleotides) of MW < 5000 Da were removed using Vivaspin columns. RNAs were oxidized at their 3'-ends using 5 mM KIO_4 for 20 min with shaking (300 rpm) at 3 °C. RNAs were then precipitated by addition of 250 mM NaOAc and 3 volumes of EtOH and cooling for 2h at -80 °C. Pellets were resolved in 100 µL 50 mM NaOAc, pH 5.0 and biotinylated with a saturated biotin hydrazine solution (1 µL) over night with shaking (300 rpm) at 4 °C. Biotinylated RNAs were then EtOH precipitated with 3 volumes of EtOH and pellets were dissolved in 50 µL Labeling Buffer (10 mM TRIS, pH 7.2, 2 mM EDTA, 1 mM TCEP). RNA yields were 1.4 µg/mL *Sc*.D1-L1-GMPS (#NTP = 418) and 1.2 µg/mL *Sc*.D135-L14-GMPS (#NTP = 638). RNA solutions were degassed and the biotinylated RNAs were labeled in the dark over night at RT by the addition of 3 aliquots (0.5 µL) of Cy5-maleimide (thiol-reactive probe, 45 mM in DMSO) after 0, 0.75 and 1.75 h at RT and shaking (300 rpm). Labeling reactions were stopped after 18 h by EtOH precipitation (250 mM NaOAc, 3 volumes EtOH) and RNA pellets were dissolved in 40 µL Labeling Buffer. Labeling efficiencies were found to be 14 % for *Sc*.D1-L1-Cy5 and 2.3 % for *Sc*.D135-L14-Cy5.

All attempts to label the carboxylic acid at the 3'-end of *Sc*.D1-L1 (transcribed and oxidized with of KIO_4 as described above) as well as a shorter RNA D1ssh (5'-GGAAUAUGCUCAACGAAAGUGAAUCAUUCGUGAGAGCUAAGUUCC-3', gift from Bernd Knobloch, oxidized with KIO_4) with Alexa Fluor 647 cadaverine (aliphatic amine, a potential carboxylic acid-reactive reagents and an Cy5-analog) in a buffer of 0.1 M CDI, 0.1 M imidazole, pH 6.0 were unsuccessful.

Polymerase chain reactions (PCR)

The PCR reactions were carried out using 0.55 µL *Pfu* polymerase (2.5 U/µL). To introduce the loop sequences, 0.3 µg initial plasmid, 1 µL primers P1 and x1 (100 µM each) or P2 and x2, and 1 µL dNTP-Mix (10 mM ATP, CTP, GTP and TTP each) were mixed with *Pfu* polymerase in 10 µL 1x *Pfu* buffer in a first PCR and the PCR products R1 or R2 were

generated using a PCR method of 25 cycles of melting for 1 min at 95 °C, annealing for 30 s at 55 °C and elongation for 2 min at 72 °C. PCR products were separated on 1 % agarose gels and subsequently extracted from the gel. In the second PCR reaction, 5 µL products R1 and R2 were added with 1 µL dNTP mix and *Pfu* polymerase in 10 µL 1x *Pfu* buffer and the PCR was conducted using the method described above to result in the desired full length plasmid inserts.



Scheme 1. Primer location.

Primer sequences

The end primers for the polymerase reactions were P1 (5'-GGGCGAATTCGTGAATT GTAATACGAC-3') and P2 (5'-CTACCCAAGCTTGCATGCCTGCAGG-3') (Scheme 1). The primer sequence x1 and the primer sequence x2 depending on the plasmid were:

pT7D135-L1: x1, 5'-CCCTACCAAGAGCGTTATTAATTTGGGATAATAATTGAAT ATCAGACAGGTTTTCAATTAGTGGTG-3', x2, 5'-CCCAAATTAATAACGCTCTTGGT AGGGATAAAATGGTTGATGTTATGTATTGGAAATGAGC-3'

pT7D135-L4: x1; 5'-CCCTACCAAGAGCGTTATTAATTTGGGTATCACCTATAGT ATAAGTTAGCAGATTTTCATC-3'; x2, 5'-CCCAAATTAATAACGCTCTTGGTAGGGT ATTGCGTGAGCCGTATGC-3'

pT7D135-L14: x1, 5'-CCCTACGTAGTCCGAAATATATTTGGGTATCACCTATAGT ATAAGTTAGCAGATTTTCATC-3'; x2, 5'-CCCAAATATATTTCGGACTACGTAGGGT ATTGCGTGAGCCGTATGC-3'

pT7D135-L1α: x1, 5'-CCCTACGTAGTCCGAAATATATTTGGGGGTAAATATTAT TTATGATAACTTTCAGACCGCTC-3'; x2, 5'-CCCAAATATATTTCGGACTACGTAGG GGGATAAATTATATTTTTATCAATATAAGTCTAATTACAAGTGTATTAAAATGG-3'

pT7D135-L1β: x1, 5'-CCCTACGTAGTCCGAAATATATTTGGGATAAAATAACTG TCAAGAATATGGATACTTTTGTCATTACAGC-3'; x2, 5'-CCCAAATATATTTCGGAC TACGTAGGGATAAAAAAAAGATGAAGGAACTTTGA CTGATC-3'

pT7D135-L1κ: x1, 5'-CCCTACGTAGTCCGAAATATATTTGGGGTTGAGCATATT AGATCAGTCAAAGTTCC-3'; x2, 5'-CCCAAATATATTTCGGACTACGTAGGGGTGA ATCAAATGTTATAAAATTACTTACACCAC-3'

Underlined nucleotides indicate inserted loop sequences. Loop sequence L1 was inserted into pT7D135 to result in pT7D135-L1 and pT7D135-L4 and loop sequence L2 was inserted

into pT7D135-L1 to result in pT7D135-L14, pT7D135-L1α, pT7D135-L1β and pT7D135-L1κ.

Primer sequences to introduce the *Eco*RI restriction site into pT7D135-L14 and to result in pT7D1-L1-ER were: x1, 5'-GATATTTACGTATTTATGATAAAACAGAATTCACCCTC TTCGG-5'; x2, 3'-CCGAAGAGGGTGAATTCTGTTTTATCATAAATACGTAAATATC-3'. Underlined positions indicate mutated nucleotides.

DNA-oligos Cy3-DNA, Cy5-DNA, T-Biotin and substrates 17/7, 17/7-dC

The biotinylated DNA strand T-Biotin (5'-Biotin-TGCATGCCTGCAGGTCGACTCTA-3'), Cy3-DNA (5-Cy3-ACCAAGAGCGTTATTAAT-3') and Cy5-DNA (5'-Cy5-ACGTAGTCCGAAATATAT-3') were purified by denaturing 18%-PAGE and subsequent C8-reversed phase HPLC chromatography as described for AP-containing RNAs.[143] The RNA-substrates 17/7 (3'-UGGCGAGCUUUUACAGGGUGGUGC-5', 2'-protected) and 17/7-dC (3'-UGGCGAGdCUUUUACAGGGUGGUGC-5', 2'-protected), unlabeled or 5'-Cy5-labeled, were deprotected as suggested by the manufacturer and subsequently also gel- and HPLC-purified.[143]

DNA-oligos T-Biotin2 and T-D1

T-Biotin2 (5'-Biotin-TGCATGCCTGCAGGTCGACTCTAX-3') and T-D1 (5'-CTTTTACCCTCCGAAGAGGGTTX-3') were ordered containing a C7-aminolinker (X). The oligonucleotides were PAGE purified as described above and subsequently, 0.3 μM DNA-oligo was labeled in a total volume of 100 μL with 14 μL Cy5-maleimide (20 M in DMSO) in 20 mM $NaCO_3$, pH 8.5. The labeling reaction was carried out at RT overnight with shaking (300 rpm). Labeled DNA-oligos were separated from unlabeled fraction and purified using C8-reversed phase HPLC chromatography as described above and stored in H_2O.

5.5.2 Single-turnover cleavage assay

The substrates 17/7, 17/7-Cy5, 17/7-dC and 17/7-dC-Cy5 were 5'-labeled with γ-^{32}P-ATP and T4 polynucleotide kinase using standard procedures or 3'-labeled by the Szostak-Splint method using the 17/7-Splint (5'-Cy5-GUGGUGGGACAUUUUCGAGCGG-3'), α-^{32}P-CTP and Klenow DNA polymerase as described for *Av*.D135.[273]

DNA-oligos were in 6 fold excess over RNAs to avoid free ribozyme in the experiments. Cy3-DNA, Cy5-DNA and T-Biotin (600 nM) were heat-annealed to *Sc*.D135-L14 (100 nM)

at 90°C for 1 min in Reaction Buffer (80 mM MOPS pH 6.9, 500 mM KCl) followed by Mg^{2+} addition and incubation at 42°C for 15 min in varying amounts of $MgCl_2$. ^{32}P-Labeled 17/7 or 17/7-dC (1 nM) was separately incubated in reaction buffer and $MgCl_2$ and then added to annealed, biotinylated and Cy3/Cy5-labeled ribozyme to start the cleavage reaction. All cleavage reactions mentioned for RNAs other than *Sc*.D135-L14 were conducted as described for *Sc*.D135-L14, only that experiments with *Sc*.D135-L1 and *Sc*.D135-L4 were missing Cy3-DNA, Cy5-DNA and T-Biotin and contained ^{32}P-labeled substrates 17/7-Cy5 and 17/7-dC-Cy5 as described in the text, figures and tables. Aliquots (1 µL) were removed from the reaction mixture (20 µL) at specific points in time and analyzed on 18% polyacrylamide gels. Reaction rate constants were determined by quantification of gels and fitting to single exponential expression:

$$frac[Product] = (1 - A_1) - A_2 \times e^{-kt} \tag{10}$$

with A_1 being the fraction of uncleaved substrate, A_2 the fraction of product formed, and k the first order rate constant in min^{-1}. In some experiments the ribozyme complex was not prefolded in $MgCl_2$, but Mg^{2+} was added only with the substrate IBS to initiate folding and cleavage reaction.

The first order rate constants were plotted in respect to increasing [Mg^{2+}] and fit to a modified Hill equation

$$y = y_0 + \frac{x^n}{K_{Mg}{}^n + x^n} \tag{11}$$

where K_{Mg} is the [Mg^{2+}] at half maximum rate and n represents the cooperativity constant.

5.5.3 Native gel assay

Samples (5 µL) containing Cy3-DNA (10 nmol), ribozyme (15 nmol) and Cy5-DNA (30 nmol or as indicated) were denatured in Reaction Buffer for 45 s at 90-95 °C. 6 mM $MgCl_2$ was added and samples were allowed to prefold for 15 min at 42 °C in a final volume of 5 µL. An equal amount of glycerol was added and the samples were loaded onto a 6 % (w/v) gel containing 3 mM MgOAc and 66 mM HEPES, 34 mM Tris, pH 7.4 (RT) as described. Elecrophoresis was performed for 1-2h at 4°C in a buffer of the same conditions as described.[278] Fluorescence bands of Cy3 and Cy5 on the gels were detected on a Typhoon molecular scanner. RNA bands were stained in an EtBr bath (5 µl EtBr/500 mL H_2O) for 15 min and destained for 5 min in H_2O for RNA band detection by UV excitation at 260 nm.

5.5.4 Steady-state FRET kinetic assay

Ratios of Cy3-DNA : ribozyme : Cy5-DNA were chosen to be 1:2.5:6 to minimize unbound amount of Cy3-DNA and maximize Cy5-binding to ribozyme for optimized fluorescence transfer. Cy3-DNA (0.1 μM), *Sc*.D135-L14 (0.25 μM) and Cy5-DNA (0.6 μM) were denatured at 90 °C for 1 min in reaction buffer and subsequently incubated at 42°C for 15-20 min. Reaction mixture (100 μL final) was excited at 555 nm and emission was recorded at 565 nm and 665 nm on a Cary Eclipse Fluorescence Spectrometer. Slits were set to 5 nm and sample was thermostatted at 42 °C. Emission time courses were measured for 15 min after addition of varying amounts of $MgCl_2$ in the range of 0-100 mM.

Normalized absolute FRET changes ($FRET_{t(15)}/FRET_{t(o)}$) were plotted against Mg^{2+} concentrations for analysis of overall Mg^{2+} dependency of folding by the Hill equation (2). Time courses of FRET ratio (defined as $E_{FRET} = I_A/(I_D + I_A)$) were fitted with a single exponential expression.

5.5.5 Anisotropy measurements

Cy3-DNA (0.15 μM) or Cy5-DNA (0.5 μM) in the presence and absence of *Sc*.D135-L14 (0.3 μM) were prepared and folded in 100 mM $MgCl_2$ as described for ssFRET measurement. Anisotropy experiments were conducted in a cuvette (100 μL) to measure fluorescence polarization using polarized filters on the spectrophotometer. Fluorescence intensities (I) of polarized excitation and emission in vertical (v, 0°) and horizontal (h, 90°) positions were recorded in all four possible combinations I_{vv}, I_{vh}, I_{hv}, I_{hh}. The anisotropy value r was calculated as described[134] by

$$r = \frac{I_{vv} - gI_{vh}}{I_{vv} + 2gI_{vh}} \text{ and } g = \frac{I_{hv}}{I_{hh}} \quad (12)$$

5.5.6 Time-resolved FRET measurements

Samples for trFRET measurements were prepared using the same Cy3-DNA:*Sc*.D135-L14*:Cy5-DNA ratios and procedure as described for ssFRET assays. Time-resolved FRET (trFRET) measurements were recorded by collecting time-resolved emission profiles of the Cy3 donor in samples (150 μL) at 0, 10. 30. 50. 70 and 100 mM up to > 40'000 peak counts using time correlated single photon counting as described.[279] Data analysis to derive distance information was performed using the time resolved fluorescence decays of sample

with donor and acceptor and one employing a donor-only RNA complex using a model of distance contributions as described.[280, 281] Since the pulsed excitation at 514 nm and the isotropic detection under magic angle conditions at 530 nm were optimized to the tetramethylrhodamine/fluorescein (T/F) FRET pair, the resulting mean distances had to be corrected by a factor of R_0(Cy3/Cy5)/R_0(T/F) = 45 Å/55 Å[281] to result in the corresponding mean distances for the Cy3/Cy5 FRET pair.

5.5.7 Single molecule FRET experiments

Cy3-DNA, Cy5-DNA and T-Biotin (10 μM each) were heat-annealed to *Sc*.D135-L14 (1 μM) in Reaction Buffer containing 0.5% mercaptoethanol. After $MgCl_2$ addition (0-100 mM final) the reaction mixture (10 μl) was incubated at 42°C for 15-20 min. The FRET-labeled, biotinylated and annealed *Sc*.D135-L14* complex was diluted to ~25-50 pM and bound to a streptavidin-coated quartz slide surface. Excess Cy3-DNA, Cy5-DNA and T-Biotin oligos were removed from the slide via a washing step with Reaction Buffer. An oxygen scavenging system consisting of 10% (wt/v) glucose, 2% (v/v) 2-mercaptoethanol, 50 μg/mL glucose oxidase, and 10 μg/mL catalase was used in all experiments to reduce photobleaching. All smFRET experiments were carried out at room temperature according to previously described procedures.[261, 282, 283]

The donor (I_D) and acceptor (I_A) fluorescence signals of optically resolved single molecules were recorded on a total internal reflection fluorescence microscope as described.[261, 283] Donor and acceptor signals were found to bleach in a single step, confirming the presence of single molecules. Single molecule traces showing dynamics and before photobleaching of typically more than 150 molecules at each concentration were manually selected and E_{FRET} values for individual Mg^{2+} concentrations were accumulated in histograms. Histogram distributions were analyzed with a double or triple gaussian equation to reveal reoccurring mean FRET values. Dwell times from single molecule trajectories were calculated and plotted in dwell time histograms for calculation of rate constants as described.[282] Control experiments in the absence of *Sc*.D135-L14 or T-Biotin did not exhibit any conformational dynamics, ruling out non-specific interactions between the label DNA or the RNA and the quartz surface. All RNA systems studied to date exhibit only negligible perturbations from surface immobilization.[260]

6. Bibliography

1. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci USA, 1990. **87**(12): p. 4576-9.
2. Caspersson, T. and J. Schultz, *Ribonucleic acids in both nucleus and cytoplasm, and the function of the nucleolus.* Proc Natl Acad Sci USA, 1940. **26**(8): p. 507-15.
3. Crick, F.H., *On protein synthesis.* Symp Soc Exp Biol, 1958. **12**: p. 138-63.
4. Nissen, P., et al., *The structural basis of ribosome activity in peptide bond synthesis.* Science, 2000. **289**(5481): p. 920-30.
5. Valadkhan, S. and J.L. Manley, *Splicing-related catalysis by protein-free snRNAs.* Nature, 2001. **413**(6857): p. 701-7.
6. Moss, E.G., *RNA interference: it's a small RNA world.* Curr Biol, 2001. **11**(19): p. R772-5.
7. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. **293**(5531): p. 834-8.
8. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.
9. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.* Nature, 2000. **403**(6772): p. 901-6.
10. Nahvi, A., et al., *Genetic control by a metabolite binding mRNA.* Chem Biol, 2002. **9**(9): p. 1043.
11. Winkler, W., A. Nahvi, and R.R. Breaker, *Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression.* Nature, 2002. **419**(6910): p. 952-6.
12. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme.* Cell, 1983. **35**(3 Pt 2): p. 849-57.
13. Kruger, K., et al., *Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena.* Cell, 1982. **31**(1): p. 147-57.
14. Cech, T.R., *Ribozymes, the first 20 years.* Biochem Soc Trans, 2002. **30**(Pt 6): p. 1162-6.
15. Doudna, J.A. and T.R. Cech, *The chemical repertoire of natural ribozymes.* Nature, 2002. **418**(6894): p. 222-8.
16. Noller, H.F., V. Hoffarth, and L. Zimniak, *Unusual resistance of peptidyl transferase to protein extraction procedures.* Science, 1992. **256**(5062): p. 1416-9.
17. Piccirilli, J.A., et al., *Aminoacyl esterase activity of the Tetrahymena ribozyme.* Science, 1992. **256**(5062): p. 1420-4.
18. Steitz, T.A. and P.B. Moore, *RNA, the first macromolecular catalyst: the ribosome is a ribozyme.* Trends Biochem Sci, 2003. **28**(8): p. 411-8.
19. Tarasow, T.M., S.L. Tarasow, and B.E. Eaton, *RNA-catalysed carbon-carbon bond formation.* Nature, 1997. **389**(6646): p. 54-7.
20. Wilson, D.S. and J.W. Szostak, *In vitro selection of functional nucleic acids.* Annu Rev Biochem, 1999. **68**: p. 611-47.
21. Gilbert, W., *The origin of life: The RNA world.* Nature, 1986. **319**: p. 618.
22. Murray, J.M. and J.A. Doudna, *Creative catalysis: pieces of the RNA world jigsaw.* Trends Biochem Sci, 2001. **26**(12): p. 699-701.
23. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.
24. Voet, D. and J.G. Voet, *Biochemistry*. 2 ed. 1995, New York: John Wiley & Sons, Inc.

25. Erat, M.C., *Two domains of branching and catalysis act as specific metal ion binding platforms within the group II intron ribozyme core*, in *Institute of Inorganic Chemistry*. 2007, University of Zürich: Zürich, Switzerland.
26. Das, R., et al., *The fastest global events in RNA folding: electrostatic relaxation and tertiary collapse of the Tetrahymena ribozyme.* J Mol Biol, 2003. **332**(2): p. 311-9.
27. Woodson, S.A., *Metal ions and RNA folding: a highly charged topic with a dynamic future.* Curr Opin Chem Biol, 2005. **9**(2): p. 104-9.
28. Shiman, R. and D.E. Draper, *Stabilization of RNA tertiary structure by monovalent cations.* J Mol Biol, 2000. **302**(1): p. 79-91.
29. Batey, R.T., R.P. Rambo, and J.A. Doudna, *Tertiary motifs in RNA structure and folding.* Angew Chem Int Ed Engl, 1999. **38**(16): p. 2326-2343.
30. Costa, M. and F. Michel, *Frequent use of the same tertiary motif by self-folding RNAs.* EMBO J, 1995. **14**(6): p. 1276-85.
31. Basu, S., et al., *A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor.* Nat Struct Biol, 1998. **5**(11): p. 986-92.
32. Shcherbakova, I., et al., *Monovalent ion-mediated folding of the Tetrahymena thermophila ribozyme.* J Mol Biol, 2004. **342**(5): p. 1431-42.
33. Christian, E.L. and M. Yarus, *Metal coordination sites that contribute to structure and catalysis in the group I intron from Tetrahymena.* Biochemistry, 1993. **32**(17): p. 4475-80.
34. Piccirilli, J.A., et al., *Metal ion catalysis in the Tetrahymena ribozyme reaction.* Nature, 1993. **361**(6407): p. 85-8.
35. Nakano, S., D.M. Chadalavada, and P.C. Bevilacqua, *General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme.* Science, 2000. **287**(5457): p. 1493-7.
36. Bertini, I., A. Sigel, and H. Sigel, *Handbook on Metalloproteins*. 2001, NY-Basel: Marcel Dekker Inc. 1108.
37. Murray, J.B., et al., *The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone.* Chem Biol, 1998. **5**(10): p. 587-95.
38. Sigel, R.K. and A.M. Pyle, *Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry.* Chem Rev, 2007. **107**(1): p. 97-113.
39. Misra, V.K. and D.E. Draper, *A thermodynamic framework for Mg^{2+} binding to RNA.* Proc Natl Acad Sci USA, 2001. **98**(22): p. 12456-61.
40. Misra, V.K. and D.E. Draper, *On the role of magnesium ions in RNA stability.* Biopolymers, 1998. **48**(2-3): p. 113-35.
41. Freisinger, E. and R.K. Sigel, *From nucleotides to ribozymes - A comparison of their metal ion binding properties.* Coord. Chem. Rev, 2007. **251**: p. 1834-1851.
42. Pyle, A.M., *Ribozymes: a distinct class of metalloenzymes.* Science, 1993. **261**(5122): p. 709-14.
43. Pyle, A.M., *Metal ions in the structure and function of RNA.* J Biol Inorg Chem, 2002. **7**(7-8): p. 679-90.
44. Sigel, R.K.O., *Group II intron Ribozymes and metal ions - A delicate relationship.* Eur Inorg Chem, 2005. **12**: p. 2281-2292.
45. Narlikar, G.J. and D. Herschlag, *Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes.* Annu Rev Biochem, 1997. **66**: p. 19-59.
46. Sigel, R.K., *Intimate relationships between metal ions and nucleic acids.* Angew Chem Int Ed Engl, 2007. **46**(5): p. 654-656.
47. Roth, A., et al., *Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions.* RNA, 2006. **12**(4): p. 607-19.

48. Heus, H.A. and A. Pardi, *Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops.* Science, 1991. **253**(5016): p. 191-4.
49. Sigel, R.K., et al., *Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif.* Nat Struct Mol Biol, 2004. **11**(2): p. 187-92.
50. Sun, X., Q. Zhang, and H.M. Al-Hashimi, *Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg^{2+} binding: role in the kissing-duplex structural transition.* Nucleic Acids Res, 2007. **35**(5): p. 1698-713.
51. Staple, D.W. and S.E. Butcher, *Pseudoknots: RNA structures with diverse functions.* PLoS Biol, 2005. **3**(6): p. e213.
52. Dill, K.A. and H.S. Chan, *From Levinthal to pathways to funnels.* Nat Struct Biol, 1997. **4**(1): p. 10-9.
53. Tinoco, I., Jr. and C. Bustamante, *How RNA folds.* J Mol Biol, 1999. **293**(2): p. 271-81.
54. Wu, M. and I. Tinoco, Jr., *RNA folding causes secondary structure rearrangement.* Proc Natl Acad Sci USA, 1998. **95**(20): p. 11555-60.
55. Buchmueller, K.L., et al., *A collapsed non-native RNA folding state.* Nat Struct Biol, 2000. **7**(5): p. 362-6.
56. Waldsich, C. and A.M. Pyle, *A folding control element for tertiary collapse of a group II intron ribozyme.* Nat Struct Mol Biol, 2007. **14**(1): p. 37-44.
57. Russell, R., et al., *Exploring the folding landscape of a structured RNA.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(1): p. 155-160.
58. Wolynes, P.G., *Recent successes of the energy landscape theory of protein folding and function.* Q Rev Biophys, 2005. **38**(4): p. 405-10.
59. Treiber, D.K. and J.R. Williamson, *Exposing the kinetic traps in RNA folding.* Curr Opin Struct Biol, 1999. **9**(3): p. 339-45.
60. Woodson, S.A., *Recent insights on RNA folding mechanisms from catalytic RNA.* Cell Mol Life Sci, 2000. **57**(5): p. 796-808.
61. Pyle, A.M., *Role of metal ions in ribozymes.* Met Ions Biol Syst, 1996. **32**: p. 479-520.
62. Lafontaine, D.A., D.G. Norman, and D.M. Lilley, *Folding and catalysis by the VS ribozyme.* Biochimie, 2002. **84**(9): p. 889-96.
63. Lilley, D.M., *Structure, folding and catalysis of the small nucleolytic ribozymes.* Curr Opin Struct Biol, 1999. **9**(3): p. 330-8.
64. Walter, N.G., et al., *In the fluorescent spotlight: global and local conformational changes of small catalytic RNAs.* Biopolymers, 2001. **61**(3): p. 224-42.
65. Murray, J.B., et al., *Capture and visualization of a catalytic RNA enzyme-product complex using crystal lattice trapping and X-ray holographic reconstruction.* Mol Cell, 2000. **5**(2): p. 279-87.
66. Ferre-D'Amare, A.R., K. Zhou, and J.A. Doudna, *Crystal structure of a hepatitis delta virus ribozyme.* Nature, 1998. **395**(6702): p. 567-74.
67. Cai, Z. and I. Tinoco, Jr., *Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite.* Biochemistry, 1996. **35**(19): p. 6026-36.
68. Ravindranathan, S., S.E. Butcher, and J. Feigon, *Adenine protonation in domain B of the hairpin ribozyme.* Biochemistry, 2000. **39**(51): p. 16026-32.
69. Butcher, S.E., F.H. Allain, and J. Feigon, *Solution structure of the loop B domain from the hairpin ribozyme.* Nat Struct Biol, 1999. **6**(3): p. 212-6.
70. Ferre-D'amare A, R. and P.B. Rupert, *The hairpin ribozyme: from crystal structure to function.* Biochem Soc Trans, 2002. **30**(Pt 6): p. 1105-9.
71. Walter, F., A.I. Murchie, and D.M. Lilley, *Folding of the four-way RNA junction of the hairpin ribozyme.* Biochemistry, 1998. **37**(50): p. 17629-36.

72. Walter, N.G., et al., *Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer.* EMBO J, 1998. **17**(8): p. 2378-91.
73. Bassi, G.S., et al., *Ion-induced folding of the hammerhead ribozyme: a fluorescence resonance energy transfer study.* EMBO J, 1997. **16**(24): p. 7481-9.
74. Wadkins, T.S., et al., *A nested double pseudoknot is required for self-cleavage activity of both the genomic and antigenomic hepatitis delta virus ribozymes.* RNA, 1999. **5**(6): p. 720-7.
75. Harris, D.A., D. Rueda, and N.G. Walter, *Local conformational changes in the catalytic core of the trans-acting hepatitis delta virus ribozyme accompany catalysis.* Biochemistry, 2002. **41**(40): p. 12051-61.
76. Butcher, S.E., *Structure and function of the small ribozymes.* Curr Opin Struct Biol, 2001. **11**(3): p. 315-20.
77. Winkler, W.C., et al., *Control of gene expression by a natural metabolite-responsive ribozyme.* Nature, 2004. **428**(6980): p. 281-6.
78. Klein, D.J. and A.R. Ferre-D'Amare, *Structural basis of glmS ribozyme activation by glucosamine-6-phosphate.* Science, 2006. **313**(5794): p. 1752-6.
79. Frank, D.N. and N.R. Pace, *Ribonuclease P: unity and diversity in a tRNA processing ribozyme.* Annu Rev Biochem, 1998. **67**: p. 153-80.
80. Morl, M. and A. Marchfelder, *The final cut. The importance of tRNA 3'-processing.* EMBO Rep, 2001. **2**(1): p. 17-20.
81. Walter, N.G. and D.R. Engelke, *Ribozymes: Catalytic RNAs that cut things, make things, and do odd and useful jobs.* The Biologist, 2002. **49**: p. 199-203.
82. Cech, T.R., *Self-splicing of group I introns.* Annu Rev Biochem, 1990. **59**: p. 543-68.
83. Jacquier, A., *Group II introns: elaborate ribozymes.* Biochimie, 1996. **78**(6): p. 474-87.
84. Pyle, A.M., O. Fedorova, and C. Waldsich, *Folding of group II introns: a model system for large, multidomain RNAs?* Trends Biochem Sci, 2007. **32**(3): p. 138-45.
85. Russell, R., et al., *Rapid compaction during RNA folding.* Proc Natl Acad Sci USA, 2002. **99**(7): p. 4266-71.
86. Woodson, S.A., *Folding mechanisms of group I ribozymes: role of stability and contact order.* Biochem Soc Trans, 2002. **30**(Pt 6): p. 1166-9.
87. Fang, X.W., et al., *The rate-limiting step in the folding of a large ribozyme without kinetic traps.* Proc Natl Acad Sci USA, 2002. **99**(13): p. 8518-23.
88. Sosnick, T.R. and T. Pan, *RNA folding: models and perspectives.* Curr Opin Struct Biol, 2003. **13**(3): p. 309-16.
89. Stahley, M.R. and S.A. Strobel, *RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis.* Curr Opin Struct Biol, 2006. **16**(3): p. 319-26.
90. Stahley, M.R. and S.A. Strobel, *Structural evidence for a two-metal-ion mechanism of group I intron splicing.* Science, 2005. **309**(5740): p. 1587-90.
91. Rangan, P. and S.A. Woodson, *Structural requirement for Mg^{2+} binding in the group I intron core.* J Mol Biol, 2003. **329**(2): p. 229-38.
92. DeRose, V.J., *Metal ion binding to catalytic RNA molecules.* Curr Opin Struct Biol, 2003. **13**(3): p. 317-24.
93. Steitz, T.A. and J.A. Steitz, *A general two-metal-ion mechanism for catalytic RNA.* Proc Natl Acad Sci USA, 1993. **90**(14): p. 6498-502.
94. Lehmann, K. and U. Schmidt, *Group II introns: structure and catalytic versatility of large natural ribozymes.* Crit Rev Biochem Mol Biol, 2003. **38**(3): p. 249-303.
95. Michel, F. and J.L. Ferat, *Structure and activities of group II introns.* Annu Rev Biochem, 1995. **64**: p. 435-61.

96. Jarrell, K.A., et al., *Group II intron self-splicing. Alternative reaction conditions yield novel products.* J Biol Chem, 1988. **263**(7): p. 3432-9.
97. Peebles, C.L., et al., *Group II intron self-splicing: development of alternative reaction conditions and identification of a predicted intermediate.* Cold Spring Harb Symp Quant Biol, 1987. **52**: p. 223-32.
98. Padgett, R.A., et al., *The stereochemical course of group II intron self-splicing.* Science, 1994. **266**(5191): p. 1685-8.
99. Podar, M., P.S. Perlman, and R.A. Padgett, *Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions.* Mol Cell Biol, 1995. **15**(8): p. 4466-78.
100. Chin, K. and A.M. Pyle, *Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection.* RNA, 1995. **1**(4): p. 391-406.
101. Daniels, D.L., W.J. Michels, Jr., and A.M. Pyle, *Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products.* J Mol Biol, 1996. **256**(1): p. 31-49.
102. Podar, M., et al., *Group II intron splicing in vivo by first-step hydrolysis.* Nature, 1998. **391**(6670): p. 915-8.
103. Vogel, J. and T. Borner, *Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing.* EMBO J, 2002. **21**(14): p. 3794-803.
104. Costa, M., F. Michel, and E. Westhof, *A three-dimensional perspective on exon binding by a group II self-splicing intron.* EMBO J, 2000. **19**(18): p. 5007-18.
105. Jacquier, A. and N. Jacquesson-Breuleux, *Splice site selection and role of the lariat in a group II intron.* J Mol Biol, 1991. **219**(3): p. 415-28.
106. Su, L.J., et al., *Guiding ribozyme cleavage through motif recognition: the mechanism of cleavage site selection by a group ii intron ribozyme.* J Mol Biol, 2001. **306**(4): p. 655-68.
107. Peebles, C.L., et al., *A self-splicing RNA excises an intron lariat.* Cell, 1986. **44**(2): p. 213-23.
108. van der Veen, R., et al., *Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro.* Cell, 1986. **44**(2): p. 225-34.
109. Qin, P.Z. and A.M. Pyle, *The architectural organization and mechanistic function of group II intron structural elements.* Curr Opin Struct Biol, 1998. **8**(3): p. 301-8.
110. Qin, P.Z. and A.M. Pyle, *Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg^{2+} ions in the tertiary structure.* Biochemistry, 1997. **36**(16): p. 4718-30.
111. Fedorova, O., L.J. Su, and A.M. Pyle, *Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions.* Methods, 2002. **28**(3): p. 323-35.
112. Pyle, A.M. and J.B. Green, *Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate.* Biochemistry, 1994. **33**(9): p. 2716-25.
113. Pyle, A.M. and A.M. Lambowitz, *Group II introns: ribozymes that splice RNA and invade DNA.* The RNA World (3rd edition), ed. R.F. Gesteland, T.R. Cech, and J.F. Atkins. 2006: Cold Spring Harbor Press. p. 469-506.
114. Koch, J.L., et al., *Group II introns deleted for multiple substructures retain self-splicing activity.* Mol Cell Biol, 1992. **12**(5): p. 1950-8.
115. Chanfreau, G. and A. Jacquier, *An RNA conformational change between the two chemical steps of group II self-splicing.* EMBO J, 1996. **15**(13): p. 3466-76.
116. Costa, M., et al., *Multiple tertiary interactions involving domain II of group II self-splicing introns.* J Mol Biol, 1997. **267**(3): p. 520-36.

117. Fedorova, O. and A.M. Pyle, *Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3.* EMBO J, 2005. **24**(22): p. 3906-16.
118. Lambowitz, A.M. and S. Zimmerly, *Mobile group II introns.* Annu Rev Genet, 2004. **38**: p. 1-35.
119. Watanabe, K. and A.M. Lambowitz, *High-affinity binding site for a group II intron-encoded reverse transcriptase/maturase within a stem-loop structure in the intron RNA.* RNA, 2004. **10**(9): p. 1433-43.
120. Ostersetzer, O., et al., *CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains.* Plant Cell, 2005. **17**(1): p. 241-55.
121. Abramovitz, D.L., R.A. Friedman, and A.M. Pyle, *Catalytic role of 2'-hydroxyl groups within a group II intron active site.* Science, 1996. **271**(5254): p. 1410-3.
122. Zhang, L. and J.A. Doudna, *Structural insights into group II intron catalysis and branch-site selection.* Science, 2002. **295**(5562): p. 2084-8.
123. Hamill, S. and A.M. Pyle, *The receptor for branch-site docking within a group II intron active site.* Mol Cell, 2006. **23**(6): p. 831-40.
124. Erat, M.C., et al., *Solution structure of domain 6 from a self-splicing group II intron ribozyme: A Mg^{2+} binding site is located close to the stacked branch adenosine.* Chembiochem, 2007. **8**(3): p. 306-314.
125. Swisher, J.F., et al., *Productive folding to the native state by a group II intron ribozyme.* J Mol Biol, 2002. **315**(3): p. 297-310.
126. Griffin, E.A., Jr., et al., *Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups.* Chem Biol, 1995. **2**(11): p. 761-70.
127. Michels, W.J., Jr. and A.M. Pyle, *Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships.* Biochemistry, 1995. **34**(9): p. 2965-77.
128. Xiang, Q., et al., *Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets.* Biochemistry, 1998. **37**(11): p. 3839-49.
129. Swisher, J., et al., *Visualizing the solvent-inaccessible core of a group II intron ribozyme.* EMBO J, 2001. **20**(8): p. 2051-61.
130. Su, L.J., M. Brenowitz, and A.M. Pyle, *An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme.* J Mol Biol, 2003. **334**(4): p. 639-52.
131. Fang, X.W., T. Pan, and T.R. Sosnick, *Mg^{2+}-dependent folding of a large ribozyme without kinetic traps.* Nat Struct Biol, 1999. **6**(12): p. 1091-5.
132. Su, L.J., C. Waldsich, and A.M. Pyle, *An obligate intermediate along the slow folding pathway of a group II intron ribozyme.* Nucleic Acids Res, 2005. **33**(21): p. 6674-87.
133. Fedorova, O., C. Waldsich, and A.M. Pyle, *Group II intron folding under near-physiological conditions: collapsing to the near-native state.* J Mol Biol, 2007. **366**(4): p. 1099-114.
134. Lakowicz, J.R., *Principles in Fluorescence Spectroscopy*. 3rd ed. 2006.
135. Rist, M.J. and J.P. Marino, *Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions.* Current Organic Chemistry, 2002. **6**(9): p. 775-793.
136. Ward, D.C., E. Reich, and L. Stryer, *Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives.* J Biol Chem, 1969. **244**(5): p. 1228-37.

137. Jean, J.M. and K.B. Hall, *2-Aminopurine fluorescence quenching and lifetimes: role of base stacking.* Proc Natl Acad Sci USA, 2001. **98**(1): p. 37-41.
138. Stryer, L., *Fluorescence energy transfer as a spectroscopic ruler.* Annu Rev Biochem, 1978. **47**: p. 819-46.
139. Stryer, L. and R.P. Haugland, *Energy transfer: a spectroscopic ruler.* Proc Natl Acad Sci USA, 1967. **58**(2): p. 719-26.
140. Bokinsky, G., et al., *Single-molecule transition-state analysis of RNA folding.* Proc Natl Acad Sci USA, 2003. **100**(16): p. 9302-7.
141. Zhuang, X.W., et al., *A single-molecule study of RNA catalysis and folding.* Science, 2000. **288**(5473): p. 2048-+.
142. *The Handbook: A guide to fluorescent probes and labeling technologies*: Molecular Probes (Invitrogen).
143. Walter, N.G., *Structural dynamics of catalytic RNA highlighted by fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 19-30.
144. Perkins, T.T., D.E. Smith, and S. Chu, *Single polymer dynamics in an elongational flow.* Science, 1997. **276**(5321): p. 2016-21.
145. Neher, E. and B. Sakmann, *Single-channel currents recorded from membrane of denervated frog muscle fibres.* Nature, 1976. **260**(5554): p. 799-802.
146. Lu, H.P., L. Xun, and X.S. Xie, *Single-molecule enzymatic dynamics.* Science, 1998. **282**(5395): p. 1877-82.
147. Kellermayer, M.S., et al., *Folding-unfolding transitions in single titin molecules characterized with laser tweezers.* Science, 1997. **276**(5315): p. 1112-6.
148. Rueda, D. and N.G. Walter, *Single molecule fluorescence control for nanotechnology.* J Nanosci Nanotechnol, 2005. **5**(12): p. 1990-2000.
149. Michel, F., A. Jacquier, and B. Dujon, *Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure.* Biochimie, 1982. **64**(10): p. 867-81.
150. Ferat, J.L. and F. Michel, *Group II self-splicing introns in bacteria.* Nature, 1993. **364**(6435): p. 358-61.
151. Martinez-Abarca, F. and N. Toro, *Group II introns in the bacterial world.* Mol Microbiol, 2000. **38**(5): p. 917-26.
152. Dai, L. and S. Zimmerly, *Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior.* Nucleic Acids Res, 2002. **30**(5): p. 1091-102.
153. Ferat, J.L., M. Le Gouar, and F. Michel, *Multiple group II self-splicing introns in mobile DNA from Escherichia coli.* C R Acad Sci III, 1994. **317**(2): p. 141-8.
154. Michel, F., K. Umesono, and H. Ozeki, *Comparative and functional anatomy of group II catalytic introns - a review.* Gene, 1989. **82**: p. 5-30.
155. Ferat, J.L., M. Le Gouar, and F. Michel, *A group II intron has invaded the genus Azotobacter and is inserted within the termination codon of the essential groEL gene.* Mol Microbiol, 2003. **49**(5): p. 1407-23.
156. Fedorova, O. and N. Zingler, *Group II introns: structure, folding and splicing mechanism.* Biol Chem, 2007. **388**(7): p. 665-78.
157. Hebbar, S.K., S.M. Belcher, and P.S. Perlman, *A maturase-encoding group IIA intron of yeast mitochondria self-splices in vitro.* Nucleic Acids Res, 1992. **20**(7): p. 1747-54.
158. Adamidi, C., O. Fedorova, and A.M. Pyle, *A group II intron inserted into a bacterial heat-shock operon shows autocatalytic activity and unusual thermostability.* Biochemistry, 2003. **42**(12): p. 3409-18.
159. Chee, G.J. and H. Takami, *Housekeeping recA gene interrupted by group II intron in the thermophilic Geobacillus kaustophilus.* Gene, 2005. **363**: p. 211-20.

160. Costa, M., et al., *A group II self-splicing intron from the brown alga Pylaiella littoralis is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation.* J Mol Biol, 1997. **274**(3): p. 353-64.
161. Fang, X.W., et al., *The thermodynamic origin of the stability of a thermophilic ribozyme.* Proc Natl Acad Sci USA, 2001. **98**(8): p. 4355-60.
162. Juneau, K. and T.R. Cech, *In vitro selection of RNAs with increased tertiary structure stability.* RNA, 1999. **5**(8): p. 1119-29.
163. Juneau, K., et al., *Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA--solvent interactions.* Structure, 2001. **9**(3): p. 221-31.
164. Brown, R.S., et al., *Pb(II)-catalysed cleavage of the sugar-phosphate backbone of yeast tRNAPhe--implications for lead toxicity and self-splicing RNA.* Nature, 1983. **303**(5917): p. 543-6.
165. Ciesiolka, J., T. Marciniec, and W. Krzyzosiak, *Probing the environment of lanthanide binding sites in yeast tRNA(Phe) by specific metal-ion-promoted cleavages.* Eur J Biochem, 1989. **182**(2): p. 445-50.
166. Berens, C., et al., *Visualizing metal-ion-binding sites in group I introns by iron(II)-mediated Fenton reactions.* Chem Biol, 1998. **5**(3): p. 163-75.
167. Boudvillain, M. and A.M. Pyle, *Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis.* EMBO J, 1998. **17**(23): p. 7091-104.
168. Dorner, S. and A. Barta, *Probing ribosome structure by europium-induced RNA cleavage.* Biol Chem, 1999. **380**(2): p. 243-51.
169. Gast, F.U., et al., *Secondary structure probing of potato spindle tuber viroid (PSTVd) and sequence comparison with other small pathogenic RNA replicons provides evidence for central non-canonical base-pairs, large A-rich loops, and a terminal branch.* J Mol Biol, 1996. **262**(5): p. 652-70.
170. Harris, D.A. and N.G. Walter, *Probing RNA structure and metal-binding sites using terbium footprinting.* Curr Protocols Nucl Acid Chem, 2003. **6.8**: p. 6.8.1-6.8.8.
171. Walter, N.G., N. Yang, and J.M. Burke, *Probing non-selective cation binding in the hairpin ribozyme with Tb(III).* J Mol Biol, 2000. **298**(3): p. 539-55.
172. Sigel, R.K. and A.M. Pyle, *Lanthanide ions as probes for metal ions in the structure and catalytic mechanism of ribozymes.* Met Ions Biol Syst, 2003. **40**: p. 477-512.
173. Shannon, R.D., *Revised effective ionic radii and systematic studies of interatomic distances in halides and chlcogenides.* Acta Crystallographica, 1976. **A32**: p. 751-767.
174. Nieboer, E., Struct and Bond, 1975. **22**: p. 1-47.
175. Sigel, R.K., A. Vaidya, and A.M. Pyle, *Metal ion binding sites in a group II intron core.* Nat Struct Biol, 2000. **7**(12): p. 1111-6.
176. Matsumura, K. and M. Komiyama, *Enormously fast RNA hydrolysis by lanthanide(III) ions under physiological conditions: eminent candidates for novel tools of biotechnology.* J Biochem, 1997. **122**(2): p. 387-94.
177. Baes, C.F., Jr. and R.E. Mesmer, *The hydrolysis of cations.* (Wiley Interscience, New York), 1976: p. 192-146.
178. Feig, A.L., W.G. Scott, and O.C. Uhlenbeck, *Inhibition of the hammerhead ribozyme cleavage reaction by site-specific binding of Tb.* Science, 1998. **279**(5347): p. 81-4.
179. Horrocks, W.D., Jr., *Luminescence spectroscopy.* Methods Enzymol, 1993. **226**: p. 495-538.
180. Harris, D.A., R.A. Tinsley, and N.G. Walter, *Terbium-mediated footprinting probes a catalytic conformational switch in the antigenomic hepatitis delta virus ribozyme.* J Mol Biol, 2004. **341**(2): p. 389-403.

181. Nwe, K., J.P. Richard, and J.R. Morrow, *Direct excitation luminescence spectroscopy of Eu(III) complexes of 1,4,7-tris(carbamoylmethyl)-1,4,7,10-tetraazacyclododecane derivatives and kinetic studies of their catalytic cleavage of an RNA analog.* Dalton Trans, 2007: p. 5171-5178.
182. Chanfreau, G. and A. Jacquier, *Catalytic site components common to both splicing steps of a group II intron.* Science, 1994. **266**(5189): p. 1383-7.
183. Peebles, C.L., et al., *Catalytically critical nucleotide in domain 5 of a group II intron.* Proc Natl Acad Sci USA, 1995. **92**(10): p. 4422-6.
184. Konforti, B.B., et al., *Ribozyme catalysis from the major groove of group II intron domain 5.* Mol Cell, 1998. **1**(3): p. 433-41.
185. Boudvillain, M., A. de Lencastre, and A.M. Pyle, *A tertiary interaction that links active-site domains to the 5' splice site of a group II intron.* Nature, 2000. **406**(6793): p. 315-8.
186. Fedorova, O., T. Mitros, and A.M. Pyle, *Domains 2 and 3 interact to form critical elements of the group II intron active site.* J Mol Biol, 2003. **330**(2): p. 197-209.
187. Jacquier, A., *Self-splicing group II and nuclear pre-mRNA introns: how similar are they?* Trends Biochem Sci, 1990. **15**(9): p. 351-4.
188. Jacquier, A. and F. Michel, *Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns.* J Mol Biol, 1990. **213**(3): p. 437-47.
189. Sontheimer, E.J., P.M. Gordon, and J.A. Piccirilli, *Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome.* Genes Dev, 1999. **13**(13): p. 1729-41.
190. Mikheeva, S., et al., *Deletion of a conserved dinucleotide inhibits the second step of group II intron splicing.* RNA, 2000. **6**(11): p. 1509-15.
191. de Lencastre, A., S. Hamill, and A.M. Pyle, *A single active-site region for a group II intron.* Nat Struct Mol Biol, 2005. **12**(7): p. 626-7.
192. Jacquier, A. and F. Michel, *Multiple exon-binding sites in class II self-splicing introns.* Cell, 1987. **50**(1): p. 17-29.
193. Gordon, P.M. and J.A. Piccirilli, *Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis.* Nat Struct Biol, 2001. **8**(10): p. 893-8.
194. Correll, C.C. and K. Swinger, *Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 A resolution.* RNA, 2003. **9**(3): p. 355-63.
195. Chu, V.T., et al., *Control of branch-site choice by a group II intron.* Embo J, 2001. **20**(23): p. 6866-76.
196. Boulanger, S.C., et al., *Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron.* Mol Cell Biol, 1995. **15**(8): p. 4479-88.
197. Schmidt, U., et al., *Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing in vitro and in vivo: phenotypes and suppressor mutations.* RNA, 1996. **2**(11): p. 1161-72.
198. Seetharaman, M., et al., *Structure of a self-splicing group II intron catalytic effector domain 5: parallels with spliceosomal U6 RNA.* RNA, 2006. **12**(2): p. 235-47.
199. Sashital, D.G., et al., *U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction.* Nat Struct Mol Biol, 2004. **11**(12): p. 1237-42.
200. Berridge, M.J., M.D. Bootman, and H.L. Roderick, *Calcium signalling: dynamics, homeostasis and remodelling.* Nat Rev Mol Cell Biol, 2003. **4**(7): p. 517-29.
201. Erat, M.C. and R.K. Sigel, *Divalent metal ions tune the self-splicing reaction of the yeast mitochondrial group II intron Sc.ai5γ.* J Biol Inorg Chem, 2008.
202. Burton, A.S. and N. Lehman, *Calcium(II)-dependent catalytic activity of the Azoarcus ribozyme: testing the limits of resolution for in vitro selection.* Biochimie, 2006. **88**(7): p. 819-25.

203. Osborne, E.M., J.E. Schaak, and V.J. Derose, *Characterization of a native hammerhead ribozyme derived from schistosomes.* RNA, 2005. **11**(2): p. 187-96.
204. Roychowdhury-Saha, M. and D.H. Burke, *Extraordinary rates of transition metal ion-mediated ribozyme catalysis.* RNA, 2006. **12**(10): p. 1846-52.
205. Horton, T.E., D.R. Clardy, and V.J. DeRose, *Electron paramagnetic resonance spectroscopic measurement of Mn^{2+} binding affinities to the hammerhead ribozyme and correlation with cleavage activity.* Biochemistry, 1998. **37**(51): p. 18094-101.
206. Gonzalez, R.L., Jr. and I. Tinoco, Jr., *Identification and characterization of metal ion binding sites in RNA.* Methods Enzymol, 2001. **338**: p. 421-43.
207. Shan, S., et al., *Defining the catalytic metal ion interactions in the Tetrahymena ribozyme reaction.* Biochemistry, 2001. **40**(17): p. 5161-71.
208. Abramovitz, D.L. and A.M. Pyle, *Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction.* J Mol Biol, 1997. **266**(3): p. 493-506.
209. Tinsley, R.A. and N.G. Walter, *Pyrrolo-C as a fluorescent probe for monitoring RNA secondary structure formation.* RNA, 2006. **12**(3): p. 522-9.
210. Stivers, J.T., *2-Aminopurine fluorescence studies of base stacking interactions at abasic sites in DNA: metal-ion and base sequence effects.* Nucleic Acids Res, 1998. **26**(16): p. 3837-44.
211. Walter, N.G., et al., *A base change in the catalytic core of the hairpin ribozyme perturbs function but not domain docking.* Biochemistry, 2001. **40**(8): p. 2580-7.
212. Jean, J.M. and K.B. Hall, *2-Aminopurine electronic structure and fluorescence properties in DNA.* Biochemistry, 2002. **41**(44): p. 13152-61.
213. Rueda, D., et al., *Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance.* Proc Natl Acad Sci USA, 2004. **101**(27): p. 10066-71.
214. Walter, N.G., *Probing RNA structural dynamics and function by fluorescence energy resonance transfer (FRET)*. Curr Protocols in Nucleic Acid Chem. 2002: John Wiley & Sons. Unit 11.10.
215. Hobartner, C. and R. Micura, *Bistable secondary structures of small RNAs and their structural probing by comparative imino proton NMR spectroscopy.* J Mol Biol, 2003. **325**(3): p. 421-31.
216. Meroueh, M. and C.S. Chow, *Thermodynamics of RNA hairpins containing single internal mismatches.* Nucleic Acids Res, 1999. **27**(4): p. 1118-25.
217. Draper, D.E., Y.V. Bukhman, and T.C. Gluick, *Thermal methods for the analysis of RNA folding pathways*. Curr. Protocols in Nucleic Acid Chem. 2000: John Wiley & Sons. Unit 11.3.
218. Draper, D.E. and T.C. Gluick, *Melting studies of RNA unfolding and RNA-ligand interactions.* Methods Enzymol, 1995. **259**: p. 281-305.
219. Norberg, J. and L. Nilsson, *Temperature dependence of the stacking propensity of adenylyl-3',5'-adenosine.* J. Phys. Chem., 1995. **99**(35): p. 13053-13302.
220. Turner, D.H., et al., *Free energy increments for hydrogen bonds in nucleic acid base pairs.* J. Am. Chem. Soc., 1987. **109**(12): p. 3783-3785.
221. Bass, B.L. and T.R. Cech, *Specific interaction between the self-splicing RNA of Tetrahymena and its guanosine substrate: implications for biological catalysis by RNA.* Nature, 1984. **308**(5962): p. 820-6.
222. Fersht, A.R., et al., *Hydrogen bonding and biological specificity analysed by protein engineering.* Nature, 1985. **314**(6008): p. 235-8.
223. SantaLucia, J., Jr., R. Kierzek, and D.H. Turner, *Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin.* Science, 1992. **256**(5054): p. 217-9.

224. Puglisi, J.D., J.R. Wyatt, and I. Tinoco, Jr., *Solution conformation of an RNA hairpin loop.* Biochemistry, 1990. **29**(17): p. 4215-26.
225. Waldsich, C. and A.M. Pyle, *A kinetic intermediate that regulates proper folding of a group II intron RNA.* J Mol Biol, 2008. **375**(2): p. 572-80.
226. Clegg, R.M., *Fluorescence resonance energy transfer and nucleic acids.* Methods Enzymol, 1992. **211**: p. 353-88.
227. Clegg, R.M., *Fluorescence resonance energy transfer.* Curr Opin Biotechnol, 1995. **6**(1): p. 103-10.
228. Millar, D.P., *Fluorescence studies of DNA and RNA structure and dynamics.* Current Opinion in Structural Biology, 1996. **6**(3): p. 322-326.
229. Leonard, N.J. and G.L. Tolman, *Fluorescent nucleosides and nucleotides.* Ann N Y Acad Sci, 1975. **255**: p. 43-58.
230. Sastry, S.S. and B.M. Ross, *A direct real-time spectroscopic investigation of the mechanism of open complex formation by T7 RNA polymerase.* Biochemistry, 1996. **35**(49): p. 15715-25.
231. Yang, C.H. and D. Soll, *Studies of transfer RNA tertiary structure of singlet-singlet energy transfer.* Proc Natl Acad Sci USA, 1974. **71**(7): p. 2838-42.
232. Ha, T., et al., *Ligand-induced conformational changes observed in single RNA molecules.* Proc Natl Acad Sci USA, 1999. **96**(16): p. 9077-82.
233. Zhuang, X.W. and M. Rief, *Single-molecule folding.* Curr Op Struct Biol, 2003. **13**(1): p. 88-97.
234. Xie, Z., et al., *Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme.* Proc Natl Acad Sci USA, 2004. **101**(2): p. 534-9.
235. Qin, P.Z. and A.M. Pyle, *Site-specific labeling of RNA with fluorophores and other structural probes.* Methods, 1999. **18**(1): p. 60-70.
236. Haughland, R.P., *Handbook of fluorescent probes and research products*. 2002: Molecular Probes.
237. Tuschl, T., et al., *A three-dimensional model for the hammerhead ribozyme based on fluorescence measurements.* Science, 1994. **266**(5186): p. 785-9.
238. Ozaki, H. and L.W. McLaughlin, *The estimation of distances between specific backbone-labeled sites in DNA using fluorescence resonance energy transfer.* Nucleic Acids Res, 1992. **20**(19): p. 5205-14.
239. Gallo, S., M. Furler, and R.K.O. Sigel, *In vitro transcription and purification of RNAs of different size.* Chimia, 2005. **59**(11): p. 812-816.
240. Yisraeli, J.K. and D.A. Melton, *Synthesis of long, capped transcripts in vitro by SP6 and T7 RNA polymerases.* Methods Enzymol, 1989. **180**: p. 42-50.
241. Bergstrom, D.E., *Orthogonal base pairs continue to evolve.* Chem Biol, 2004. **11**(1): p. 18-20.
242. Vaish, N.K., et al., *Expanding the structural and functional diversity of RNA: analog uridine triphosphates as candidates for in vitro selection of nucleic acids.* Nucleic Acids Res, 2000. **28**(17): p. 3316-22.
243. Switzer, C.Y., S.E. Moroney, and S.A. Benner, *Enzymatic recognition of the base pair between isocytidine and isoguanosine.* Biochemistry, 1993. **32**(39): p. 10489-96.
244. Rueda, D., et al., *The 5' leader of precursor tRNAAsp bound to the Bacillus subtilis RNase P holoenzyme has an extended conformation.* Biochemistry, 2005. **44**(49): p. 16130-9.
245. Burgin, A.B. and N.R. Pace, *Mapping the active site of ribonuclease P RNA using a substrate containing a photoaffinity agent.* Embo J, 1990. **9**(12): p. 4111-8.
246. Chu, B.C., G.M. Wahl, and L.E. Orgel, *Derivatization of unprotected polynucleotides.* Nucleic Acids Res, 1983. **11**(18): p. 6513-29.

247. Moore, M.J. and P.A. Sharp, *Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites.* Science, 1992. **256**(5059): p. 992-7.
248. Smith, G.J., et al., *Efficient fluorescence labeling of a large RNA through oligonucleotide hybridization.* RNA, 2005. **11**(2): p. 234-9.
249. Kulzer, F. and M. Orrit, *Single-molecule optics.* Annu Rev Phys Chem, 2004. **55**: p. 585-611.
250. Ha, T., et al., *Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism.* Proc Natl Acad Sci USA, 1999. **96**(3): p. 893-8.
251. Weiss, S., *Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy.* Nat Struct Biol, 2000. **7**(9): p. 724-9.
252. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science, 1999. **283**(5408): p. 1676-83.
253. Axelrod, D., T.P. Burghardt, and N.L. Thompson, *Total internal reflection fluorescence.* Annu Rev Biophys Bioeng, 1984. **13**: p. 247-68.
254. Deniz, A.A., et al., *Ratiometric single-molecule studies of freely diffusing biomolecules.* Annu Rev Phys Chem, 2001. **52**: p. 233-53.
255. Kwan, A., J. Dudley, and E. Lantz, *Who really discovered Snell's law?* Physics World, 2002. **15**(4): p. 64.
256. Wolf, K.B., *Geometry and dynamics in refracting systems.* Eur Physics, 2005. **16**: p. 14-20.
257. Thio, T., *A bright future for subwavelength light sources.* American Scientist, 2006. **94**: p. 40-47.
258. Pljevaljcic, G., D.P. Millar, and A.A. Deniz, *Freely diffusing single hairpin ribozymes provide insights into the role of secondary structure and partially folded states in RNA folding.* Biophys J, 2004. **87**(1): p. 457-67.
259. Green, N.M., *Avidin and streptavidin.* Methods Enzymol, 1990. **184**: p. 51-67.
260. Zhuang, X., *Single-molecule RNA science.* Annu Rev Biophys Biomol Struct, 2005. **34**: p. 399-414.
261. Zhuang, X., et al., *Correlating structural dynamics and function in single ribozyme molecules.* Science, 2002. **296**(5572): p. 1473-6.
262. Hodak, J.H., et al., *Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET.* Proc Natl Acad Sci USA, 2005. **102**(30): p. 10505-10.
263. Tan, E., et al., *A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate.* Proc Natl Acad Sci USA, 2003. **100**(16): p. 9308-13.
264. Hodak, J.H., et al., *Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10505-10.
265. Khvorova, A., et al., *Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity.* Nat Struct Biol, 2003. **10**(9): p. 708-12.
266. Penedo, J.C., et al., *Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements.* RNA, 2004. **10**(5): p. 880-8.
267. Huang, H.R., et al., *The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function.* Proc Natl Acad Sci USA, 2005. **102**(1): p. 163-8.
268. Seraphin, B., et al., *Mitochondrial splicing requires a protein from a novel helicase family.* Nature, 1989. **337**(6202): p. 84-7.
269. Solem, A., N. Zingler, and A.M. Pyle, *A DEAD protein that activates intron self-splicing without unwinding RNA.* Mol Cell, 2006. **24**(4): p. 611-7.
270. Del Campo, M., et al., *Do DEAD-box proteins promote group II intron splicing without unwinding RNA?* Mol Cell, 2007. **28**(1): p. 159-66.

271. Davanloo, P., et al., *Cloning and expression of the gene for bacteriophage T7 RNA polymerase.* Proc Natl Acad Sci USA, 1984. **81**(7): p. 2035-9.
272. rueda@chem.wayne.edu.
273. Huang, Z. and J.W. Szostak, *A simple method for 3'-labeling of RNA.* Nucleic Acids Res, 1996. **24**(21): p. 4360-1.
274. Andrus, A. and R.G. Kuimelis, *Analysis and purification of synthetic nucleic acids using HPLC*. Curr. Protocols in Nucleic Acid Chem. 2000: John Wiley & Sons. Unit 10.5.
275. Sigel, R.K., E. Freisinger, and B. Lippert, *Effects of N7-methylation, N7-platination, and C8-hydroxylation of guanine on H-bond formation with cytosine: platinum coordination strengthens the Watson-Crick pair.* J Biol Inorg Chem, 2000. **5**(3): p. 287-99.
276. Erat, M.C. and R.K. Sigel, *Determination of the intrinsic affinities of multiple site-specific Mg^{2+} ions coordinated to domain 6 of a group II intron ribozyme.* Inorg Chem, 2007. **46**(26): p. 11224-34.
277. Marky, L.A. and K.J. Breslauer, *Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves.* Biopolymers, 1987. **26**(9): p. 1601-20.
278. Rangan, P., et al., *Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme.* Proc Natl Acad Sci USA, 2003. **100**(4): p. 1574-9.
279. Eis, P.S. and D.P. Millar, *Conformational distributions of a four-way DNA junction revealed by time-resolved fluorescence resonance energy transfer.* Biochemistry, 1993. **32**(50): p. 13852-60.
280. Miick, S.M., et al., *Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions.* Proc Natl Acad Sci USA, 1997. **94**(17): p. 9080-4.
281. Walter, N.G., J.M. Burke, and D.P. Millar, *Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction.* Nat Struct Biol, 1999. **6**(6): p. 544-9.
282. Rueda, D., et al., *Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance.* Proc Natl Acad Sci U S A, 2004. **101**(27): p. 10066-71.
283. Ha, T., *Single-molecule fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 78-86.

Acknowledgement

I would like to thank Prof. Roland Sigel for giving me the opportunity to work on this fascinating project in his group and introducing me to the RNA world and community. The way that you let me realize my ideas and work with techniques that were new to you and the inorganic institute at the University of Zürich, the continuous support wherever possible and the many discussions late at night at conferences and other social meetings are greatly appreciated and will be remembered forever.

Huge thanks go to Prof. David Rueda at Wayne State University (WSU) in Detroit, MI (USA). The realization of the FRET study would not have been possible without the wonderful collaboration with you and your lab. Your interest and firm believe in this specific project and your dedication to science and life as a whole deeply impress me every time again. I very much enjoyed and profited from the many discussions and the various experiences during my stays in your lab in the eclectic city of Detroit and the many phone calls and mails afterwards. I wish you all the best for your academic and personal future.

Many thanks go to the Sigel and Freisinger groups at University of Zürich. Michèle Erat was the one that tempted me away from proteins and induced me to broaden my scientific range to RNA during my Ph.D. She also read and assessed all the parts of my thesis in an incredibly short period of time and greatly helped me to advance during the years in the lab and while writing. I had a very nice time with my labmates Sofia Gallo and Daniela Kruschel with a lot of laughs and scientific and personal up and downs, as well as Veronica Zelenay and Ariel Mucha, who were sharing the lab with us for some months. Ariel, thanks for the Mad Dogs and all the other polish vodka drinks. Maya Furler and Susann Paulus, thanks for all the countless technical tips, tricks and support during every days work in the lab. Thanks also go to Bernd Knobloch for priceless chemical advices, Siran Lu for conducting the UV-thermal measurements on D5 ahead of my time, my students Severin Schneebeli and Markus Müller for their efforts, Silke Johannsen, Maximiliane Korth and all other past and present members of the Sigel and the Freisinger Groups for the nice atmosphere on our floor. Daniela and Silke, I greatly appreciate your help on getting the translation of the summary into correct german language.

Big thanks go to Krishanthi Karunatilaka, Rajan Lamichhane and Elvin Alemán for incredible help and support and the wonderful time in the lab at WSU. In addition, you introduced me to the captivating world of asian cuisine and south-american Christmas drinks (in summer...), experiences that I would not want to have missed out on. Further thanks go to Rui Zhao and Zhuojun Guo for help on the single molecule microscope, Annamalay Venkatachalam for the trFRET measurements, Oksana Valniuk for her commitment and all the other past and present members of the Rueda group at WSU for the nice time.

Furthermore, I would like to thank the Inorganic Institute at the University of Zürich and its important people: Nathalie Fichter, Patrizia Allegro and Beatrice Spichtig for secretary services, Manfred Jöhri for computer support, Hans-Peter Stalder for mechanistic help with equipment that needed to be fixed, Lukas Kromer for supervision of the Radiochemistry lab and the invaluable help around the institute as well as being a great friend, Tristan Corbière for surviving the assistantship in the GPC-Praktikum together and for being an inspiring guy, Christian Frech and Dominik Brühwiler for invaluable discussions, and the lunch group from the Alberto Group (Lukas, Henrik, Paul, Benjamin, Philipp K., Philipp A., Selvi, Patricia) for many interesting and motivating lunch times, as well as everyone else I forgot to mention.

Important thanks go to Prof. Nils G. Walter and his lab for early start up help on FRET and for permission to conduct the very first FRET experiments in his lab at University of Michigan in Ann Arbor, MI (USA), in particular Elizabeth McDowell and Tony Manzo for help on their single molecule microscope. Thanks also go to Prof. Anna M. Pyle and her group at Yale University in New Haven, CT (USA) for general support, especially Christina Waldsich and Olga Fedorova for their input in RNA purification and folding.

Very special thanks go to Dr. Christian L. Degen and Dr. Martino Poggio for showing me their versions of life in research during and after the Ph.D. time, as well as for long-range runs and for sharing a lot of fantastic sushi and pizza "recovery" meals during my time in Los Gatos, CA (USA).

Finally, I want to thank my family and best friends. Without my parents Mariann und Dölf and my siblings Stefan and Denise, simply none of this would have been possible. Christian, thanks for your continuous love and support and everything you and your family

give me. Claudia, Andrea, Stephanie, Sereina, Corinne, 2x Daniel, Hermann, thanks for being my friends. All of these people have greatly helped me to arrive at this last full stop.